Seismic Array Processing with Northern Finland Seismological Network

Henrik Jänkävaara

April 11, 2019

Master's Thesis
Geophysics
University of Oulu

Oulun yliopisto
Kaivannaisalan tiedekunta
kypsyysnäyte

TIIVISTELMÄ OPINNÄYTETYÖSTÄ (liite FM-tutkielmaan)
Maisterintutkinnon

Pääaine Geofysiikka

Tekijä (Sukunimi ja etunimet)	**Tutkielman sivumäärä**
Jänkävaara Henrik Johan	66

Työn nimi

Seismic Array Processing with Northern Finland Seismological Network

Asiasanat: maanjäristyksien paikannus, pitkäperiodinen pinta-aalto, seisminen monipisteasema, seismic array, beamforming, beampacking

Tiivistelmä

Sodankylän geofysiikan observatorio ylläpitää yhdeksästä laajakaistaseismometreillä varustetusta asemasta koostuvaa Northern Finland Seismological Network (FN) verkkoa. Neljä asemaa tarjoaa seismistä dataa datakeskuksiin ja viisi asemaa on testivaiheessa. Pisin etäisyys asemien välillä on noin 455 kilometriä ja asemat sijaitsevat epäsäännöllisessä muodostelmassa Pohjois-Suomessa. Tutkielmassa yhdeksän aseman verkkoa käytetään seismisenä monipisteasemana (eng. array). Niin kutsuttaja beamforming- ja beampacking-menetelmiä sovelletaan käyttäen verkkoa. Beamforming-menetelmä kasvattaa käytettävän datan signaalikohinasuhdetta, mikä mahdollistaa esimerkiksi eri seismisten faasien tarkemman tarkastelun. Menetelmän käyttämiseksi tarvitaan tieto tutkittavan faasin etenemissuunnasta ja näennäisestä nopeudesta, mitkä saadaan selville esimerkiksi beampacking-menetelmällä. Käyttäen verkkoa monipisteasemana, maanjäristyksien episentrien suuntien paikannuksen tarkkuutta tarkastellaan hyödyntäen valmiiksi testejä varten laadittuja maanjäristyskatalogeja. Testeissä seismisinä signaaleina käytetään pitkäperiodisia pinta-aaltoja ja episentrin suunta selvitetään beampacking-menetelmällä. Lisäksi tutkielmassa esitellään menetelmä asemaverkon parantamiseksi nimenomaan monipisteasemakäyttöä ajatellen. Uusien asemapaikkojen etsintä perustuu verkon siirtofunktion sivumaksimien minimoimiseen. Selvästi erottuvat sivumaksimit siirtofunktiossa ovat merkkejä verkon epätarkkuudesta monipisteasemakäytössä.

Lähteitä:

Kozlovskaya, E. et. al. Seismic observations at the Sodankylä Geophysical Observatory: history, present, and the future, Geoscientific Instrumentation, Methods and Data Systems, volume 5, 2016.

Schweitzer, J. et. al. Seismic Arrays, New Manual of Seismological Observatory Practice 2, volume 2, 2012.

Shearer, P. M. Introduction to Seismology, 1999.

Wang, J. A Scheme for Initial Beam Deployment for the International Monitoring System Arrays, Pure and Applied Geophysics, volume 159, 2002.

Muita tietoja

Päiväys: 11 / 4 201 9 **Laatijan allekirjoitus:** Henrik Jk

Abstract

Sodankylä Geophysical Observatory (SGO) operates and maintains a 9-station Northern Finland Seismological Network (FN) located in northern Finland. Four of the stations are providing the data to the European Integrated Data Archive (EIDA) while five stations are working in the test regime. Every station is equipped with a 3-component broadband seismometer. Seismic array processing methods are applied using the 9-station network, which as an array is irregular and relatively large with an aperture of 455 km. As a result, the array is suitable for detecting seismic signals with large wavelengths. In this thesis, earthquake localization tests utilizing beampacking and beamforming are carried out by detecting vertical components of long-period surface wave phases. Properties of the network as an array are discussed and a method to improve the network solely for array purposes is presented. The method searches for new locations for seismic stations to reduce sidelobes in the array transfer function.

keywords: seismic array, beamforming, beampacking, earthquake localization, long-period, surface waves.

Acknowledgments

The author would like to thank thesis supervisor Elena Kozlovskaya, Kari Moisio, and Janne Narkilahti for comments and suggestions regarding the thesis. In addition, the author would like to thank the staff of SGO (Hanna Silvennoinen, Janne Narkilahti, Riitta Hurskainen, and Thomas Ulich) for the access to data of the nine seismic stations operated by SGO, and for providing a proper workspace to work with the data and to write the thesis. The study was supported by the project "FIN-EPOS - a FINnish national initiative of the European Plate Observing System" funded by the Academy of Finland in 2016-2020 (Decision no. 293449).

Contents

Abbreviations

BAZ	Back azimuth
EPOS	The European Plate Observing System
FGI	Finnish Geospatial Research Institute
FIN-EPOS	FINnish national initiative of the European Plate Observing System
FN	Northern Finland Seismological Network
LTA	Long time average
NHNM	New high noise model
NLNM	New low noise model
PPSD	Probabilistic power spectral density
PSD	Power spectral density
SGO	Sodankylä Geophysical Observatory
SNR	Signal-to-noise ratio
STA	Short time average

1 Introduction

Northern Finland Seismological Network (FN) is a seismic network operated by Sodankylä Geophysical Observatory (SGO) of the University of Oulu. The core of the network is formed by four seismic stations providing data to the European Integrated Data Archive (EIDA) while five stations are working in the test regime. In this thesis, the FN network refers to the seismic network consisting of nine stations.

In this thesis, the possibilities of using the FN network with seismic array processing methods, are examined. The thesis includes some basic information about the FN network and seismic networks in general, seismic arrays, and basic theory of seismology and array methods. Properties of the FN network as an array are discussed and experiments using array methods with FN network are carried out. Most notably, the experiments include attempts to locate earthquakes using beampacking.

In addition, some aspects to consider when searching for a location for a new seismic station are introduced. The presented method attempts to minimize sidelobes in the array transfer function by finding the most suitable location in that regard.

As for seismic signals of interest, the focus is on long-period surface waves. Surface waves are used for event detection and localization for example by Global Centroid Moment Tensor Catalogue, The Global CMT Project (2018) and in the case of glacial earthquakes by Olsen and Nettles (2017). Surface waves travel longer distances than body waves, and thus, are more easily detectable all over the world. In addition, some seismic events induce only surface waves. FN network as a large aperture array is suitable for detecting long-wavelength and lond-period seismic phases such as surface waves.

1.1 Seismic networks

The operation of the first seismic networks started in the 1960s. Back then, the main feature that dissociated the networks from the single stations, was centralized data collecting and timing and real-time data transfer. Nowadays, seismic networks are defined by the field of interest and spatial coverage of the networks, which ranges from local to regional and global, rather than by the means of data handling.

One of the essential objectives of seismic networks is the determination of earthquake locations and magnitudes. Other uses of seismic networks include earthquake early warning systems, nuclear test monitoring and induced seismicity monitoring for example in mining areas.

During the last couple of decades, major improvements in data transfer and storage capabilities have made the seismic data from different network operators easily accessible and available in near real-time. Using the data provided by different datacenters, it is possible to create virtual networks by combining the data of different seismic networks. (Havskov et al., 2012)

1.2 Seismic arrays

The development of seismic arrays emerges from the need to monitor nuclear tests and the compliance of nuclear test ban treaties. Compared to single stations, seismic arrays provide a tool to detect, locate and identify seismic events with better accuracy. The first arrays were developed and deployed starting from the early 1960s. (Schweitzer et al., 2012; Mykkeltveit et al., 1990; Rost and Thomas, 2002)

The differences between the seismic arrays and networks are mainly in the means of data analysis. In addition, seismic arrays are generally smaller than networks and must

be scaled according to the seismic wavelength of interest. In theory, any seismic network can be treated as an array, which on the other hand does not mean that the network in question would be suitable for array signal processing. In the array signal processing, the network is treated as one single system. (Schweitzer et al., 2012)

The main advantage of arrays over networks is the increased signal-to-noise ratio (SNR), which is a result of the so-called beamforming process. To increase the SNR using beamforming, coherent seismic signals between the sensors are required across the array, the noise, on the other hand, must be incoherent. The increased SNR allows, for example, to detect seismic phases that would otherwise be undetected. In addition, array signal analysis provides an additional way to estimate the direction and velocity of propagation of seismic wavefront. The direction of propagation, together with seismic phase identification, can be used to locate and identify the seismic source, for example. (Schweitzer et al., 2012; Rost and Thomas, 2009)

1.3 FN network

The first seismic experiments at SGO date back to 1954 when a seismic station (station code SOD) was founded in Tähtelä, Sodankylä. After the first initiations in Sodankylä, three new sites were established during the following decades, Oulu (station code OUL) around 1959-1963, Maaselkä (station code MSF) in 1970 and Rovaniemi (station code RNF) in 2008. The current seismic station (station code SGF) in Sodankylä was established in 2001 and it is now located in the underground tunnel. The Oulu station has been in its current location in Huttukylä since 1996.

In the current modernized form, with 3-component broadband seismometers and digital data loggers, the data of four stations (SGF, OUL, MSF, and RNF) has been online and in operation since 2008, providing real-time data under the International Federation of Digital Seismograph Networks (FDSN) with network code FN.

Figure 1 shows the locations of the FN network stations. In addition, in the map are included nearby stations of other regional networks.

For a more detailed description of the history of FN network and seismic research at SGO, see Kozlovskaya et al. (2016).

1.4 The European Plate Observing System

The European Plate Observing System (EPOS) is a Europewide research infrastructure which collects and incorporates multidisciplinary solid earth geophysical data and research infrastructures provided by different national research institutes and organizations. The plan is to integrate heterogenous geophysical data into one single easily accessible platform. The data and infrastructure are focused on solid earth research and is provided as open access. EPOS aims to provide services to serve for example scientific and educational purposes. (EPOS, n.d.; Kozlovskaya et al., 2016; Korja and Vuorinen, 2016)

1.5 FIN-EPOS

FIN-EPOS (a FINnish national initiative of the EPOS) is a consortium formed by Universities of Helsinki (UH) and Oulu (UO), National Land Survey of Finland (NLS), Geological Survey of Finland (GTK), Finnish Meteorological Institute (FMI), Finnish Geospatial Research Institute (FGI), CSC - IT Center for Science Ltd. and VTT Technical Research Centre of Finland Ltd. The consortium plans to collect and standardize, to some extent, varied geophysical data provided by participating institutes. (Korja and Vuorinen, 2016)

As a part of the FIN-EPOS subproject, Finnish seismic network is being upgraded,

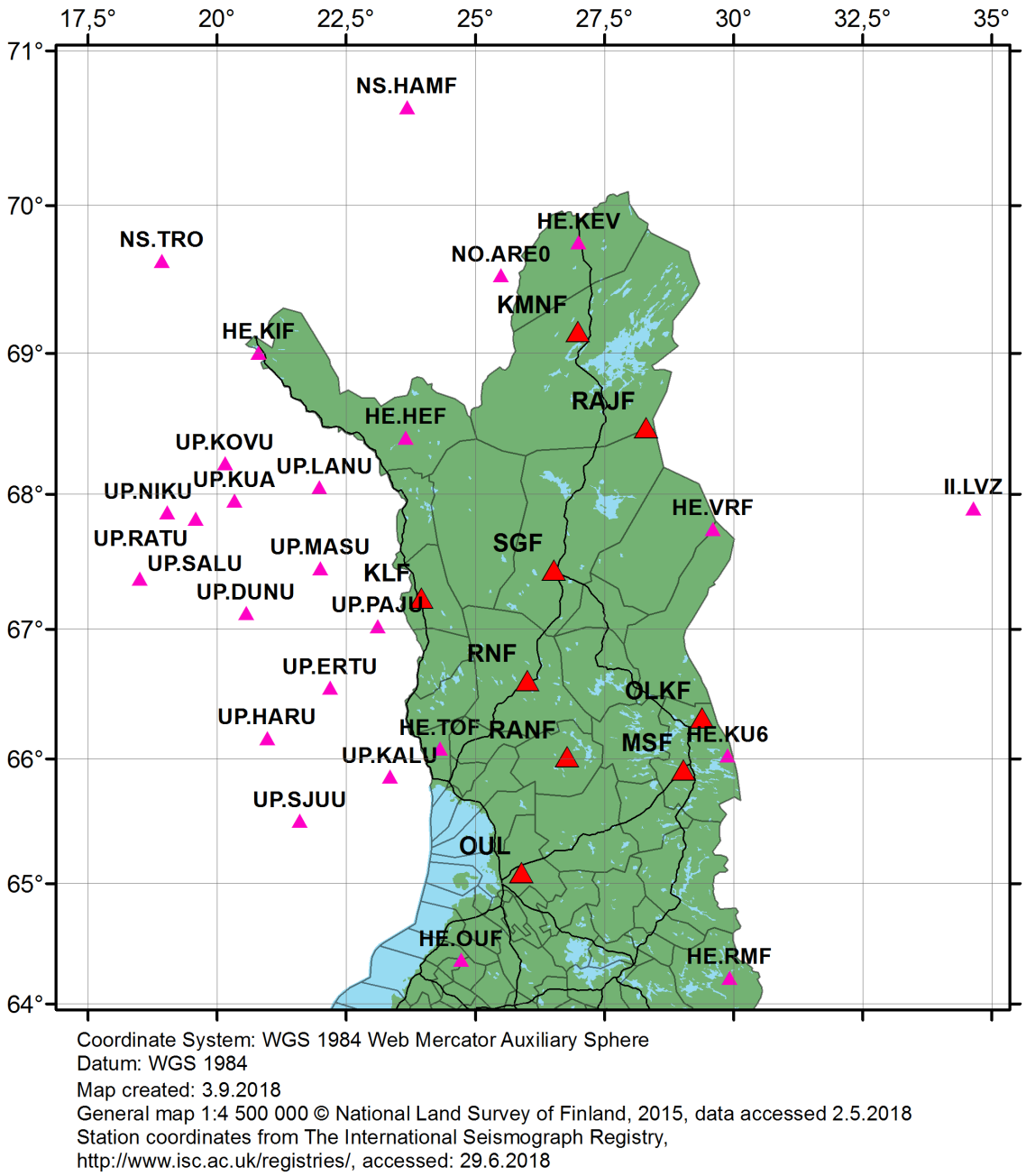


Figure 1: FN network station locations map with nearby stations of other regional networks included. The red triangles are the FN stations, and the smaller magenta triangles are stations of HE, UP, NS, NO, and II networks.

which in the case of FN network means an addition of multiple new stations. The upgrades are funded by the Academy of Finland and University of Oulu. The improvements of the FN network started in 2014 with the addition of two new stations, KLF

and OLKF. Since then three more stations have been deployed, KMNF, RANF, and RAJF. All five most recent stations are in the testing phase. (Kozlovskaya et al., 2016; Narkilahti et al., 2016)

1.6 Thesis

In this thesis, general array signal processing techniques are applied using the data acquired with the FN network which currently (as of 2018) consists of nine stations arranged in irregular formation. The purpose of the thesis is to investigate the operability of the FN network as an array and to discuss possible improvements that could be conducted in order to enhance the performance of the network as an array. These improvements can be for example new station locations chosen specifically to tune the array properties.

The locations of the current stations were selected taking into account geological conditions and ambient noise level at the site. In addition, an existing infrastructure, such as ready access to power and road accessibility were taken advantage of.

For the calculations and analysis, an array processing package was created. Python programming language, together with supporting modules was used for coding. For more information about the package and coding process, see the Appendix.

2 Theory

2.1 Seismic waves

Natural earthquakes and different human activities (explosions, mining, traffic, etc.) induce seismic waves which propagate inside the earth. The waves emerge from the focus of the seismic event and radiate across the earth. Seismic waves can be divided into body and surface waves. P-waves and S-waves are body waves which travel inside the earth and Love and Rayleigh waves are surface waves. The amplitudes of the surface waves decrease as the propagation depth inside the earth is increasing. The velocity of the waves is different. P-waves are the fastest, then S-waves and the slowest waves are surface waves. (Grotzinger and Jordan, 2014)

Let us consider three-dimensional cartesian coordinates, x-direction being the direction of the propagation of the wave. In the case of P-wave particles oscillate only along the x-axis. In that sense, P-waves are similar to acoustic waves. Acoustic and P-waves can be called compressional and they both generate volume changes in direction of propagation. S-waves, on the other hand, introduce particle motion only in yz-plane. (Shearer, 1999; Grotzinger and Jordan, 2014)

2.1.1 Surface waves

There are two types of seismic surface waves, called Love and Rayleigh waves. Let us introduce an xy-plane which is the boundary of the medium where the waves propagate, waves still traveling along the x-axis. Rayleigh waves are a result of P-wave and vertically polarized S-wave and the waves oscillate elliptically in xz-plane. Love waves, on the other hand, arise as a superposition of different reflected horizontally polarized S-wave phases. Love wave particle displacement occurs along the y-axis.

As mentioned before, the amplitude of both waves decreases with the increasing depth. The surface waves attenuate at the rate of $1/r$, where r is the traveled distance. The surface waves from strong earthquakes can circulate around the earth multiple times before attenuating out completely. Superpositions of the surface waves circulating around the earth are called normal modes or free oscillations, but we are not going to focus on those in this thesis.

Surface waves dispersion is caused by the inhomogeneous seismic velocity profile of the earth. The larger wavelengths of the surface wavefronts propagate deeper inside the earth than the shorter wavelengths. As the depth increases, the seismic velocity generally increases as well, which leads to surface waves with larger wavelengths and longer periods to propagate with higher velocities, i.e., dispersion. (Lay and Wallace, 1995; Shearer, 1999; Grotzinger and Jordan, 2014)

2.1.2 Seismic wave equations

The wave equation concerning seismic waves propagating in isotropic and homogeneous medium can be presented as follows:

$$\frac{\lambda + 2\mu}{\rho}\nabla^2\vec{u} = \frac{\partial^2\vec{u}}{\partial t^2}, \tag{2.1}$$

Where λ and μ are so-called Lamé constants, i.e., elastic constants, ρ is the density of the medium, $\vec{u}$ is the displacement of a point from its original position. In the context of elasticity, μ is called shear modulus G, i.e., the ratio of shear stress to shear strain. λ parameter's physical meaning as its own is harder to interpret.

Young's modulus, E

$$E = \frac{\mu(3\lambda + 2\mu)}{\lambda + \mu} \tag{2.2}$$

Poisson's ratio, ν

$$\nu = \frac{\lambda}{2(\lambda + \mu)} \tag{2.3}$$

Bulk modulus, K

$$K = \lambda + \frac{2\mu}{3} \tag{2.4}$$

See Eloranta (2003) and Shearer (1999) for more details.

2.2 STA/LTA trigger for detecting seismic events

In seismology, practices to find and extract seismic signals corresponding to seismic events such as earthquakes and explosions from the continuous seismic records are commonly needed. STA/LTA trigger is a simple and widely used, not very sophisticated nor advanced trigger method. It is easy to implement and use and, in this thesis, it is the trigger of choice. Of course, other means of triggering also exists.

STA/LTA is a basic triggering method used to extract seismic signals from continuous data. When using the trigger short time average (STA) and longtime averages (LTA) are calculated constantly. When the ratio (STA/LTA) of said averages exceeds a predefined value, the trigger is turned on. After the trigger is on and ratio falls under other predefined value, the trigger is turned off.

Parameters to be set in order to use STA/LTA trigger are lengths of STA and LTA and on and off thresholds. (Trnkoczy, 2012)

2.3 Probabilistic power spectral density

In order to characterize noise conditions and to estimate the overall noise levels at the seismic site, a probabilistic power spectral density (PPSD) method described by McNamara and Buland (2004) and Peterson et al. (1993), is used. Sources of continuous seismic noise at the station will be visible in PPSD plots. For example, oceanic microseisms are present in PPSD plots as a maximum around a period of 4 seconds.

To calculate PPSD plots, as described by McNamara and Buland (2004), continuous seismic data is divided into 1-hour long segments overlapping 50 percents. These segments are processed and power spectral density (PSD) of each segment is calculated. The PSDs are added to a common period-power plot to form histograms for each period. As a result, we get a probabilistic power spectral density plot of the used timeframe. From the PPSD plot, the distribution, and range of noise levels of each period can be studied.

In PPSD plots new high and low noise models (NHNM and NLNM) by Peterson et al. (1993) are shown. NLNM provides some kind of reference to the minimum seismic noise level that is to be expected to be present at any seismic site. Values of PPSDs calculated using long term seismic data should be mostly contained between NHNM and NLNM.

It is important to notice that the method does not do any distinction between seismic event and noise and in the seismic data used for the method, both events and noise are present. This means that in order to get reliable results one must use the appropriate amount of continuous seismic data so that included events are not noticeable, in other

words, the vast majority of the data consists of only noise. Since the proportion of the events in continuous seismic data is generally much lower than the proportion of the noise, increasing the used timespan for PPSD mitigates the impact of events in the calculations.

In addition, it must be noted that the PPSD method will not tell anything about the SNR at the location.

For more information about preprocessing the data for PPSD calculation and about the calculations themselves, see McNamara and Buland (2004).

2.4 Array methods

2.4.1 Array transfer function

To estimate the response of the array for certain wavenumber vectors we can calculate the array transfer function, which describes received energy by the array as a function of the difference of two wavenumbers $\mathbf{k}_0 - \mathbf{k}$. The array transfer function $|C(\mathbf{k}_0 - \mathbf{k})|^2$ is presented as follows by Rost and Thomas (2009):

$$|C(\mathbf{k}_0 - \mathbf{k})|^2 = |\frac{1}{M} \sum_{j=1}^{M} e^{2\pi i \cdot \mathbf{r}_j \cdot (\mathbf{k}_0 - \mathbf{k})}|^2 \tag{2.5}$$

where M is the number of stations and $\mathbf{r}_j$ is the location vector of each station. In this thesis, location and wavenumber vectors are 2-dimensional.

When calculating the transfer function using Equation 2.5 with a range of wavenumber differences $\mathbf{k}_0 - \mathbf{k}$, produces a relative power plot as a function certain deviations from the reference wavenumber $\mathbf{k}_0$. Examples of the resulting plots can be seen in Figure 8.

The center maximum is called mainlobe and the other maxima are sidelobes. Sidelobes in the transfer function indicate ambiguities in the detection of signals with certain wavenumber combinations. In the wavenumber space of interest, sidelobes should not be prominent and the azimuthal dependency of the sidelobes should be minimal. (Rost and Thomas, 2009)

Since the wavenumber $\mathbf{k}$ is related to the slowness $\mathbf{s}$ and frequency f as follows:

$$\mathbf{k} = 2\pi f \mathbf{s}, \tag{2.6}$$

the array resolution with different slowness and frequency combinations can be analyzed more closely, as it is done in Chapter 3.3. In seismology, quantity slowness, which is the reciprocal of velocity, is generally used.

2.4.2 Plane-wave approximation

In array methods, it is assumed that the wavefront approaching the array can be approximated by a plane. According to Schweitzer et al. (2012), the plane-wave approximation is valid if the distance to the source is more than 10 times the wavelength of the wavefront and if the wavelength is not significantly larger than the aperture of the array. Aperture is the largest dimension of the array. Assuming plane-wave, the back azimuth (BAZ) and the incidence angle of the event will be the same for all stations, see Figures 2 and 3.

In addition, the seismic source is treated as a point source, which is a valid approximation if the distance to the source is more than a few wavelengths, meaning that the source lies in the so-called far-field region. The requisite for this approximation is already fulfilled when assuming plane-wave. (Aki and Richards, 2002)

The requirements for the plane-wave approximations arranged into more mathematical form:

$$d > 10\lambda,\ a \gtrsim \lambda, \tag{2.7}$$

where d is the distance to the source, λ is the wavelength, and a is the aperture of the array.

2.4.3 Beamforming

A method to increase SNR in the array is beamforming. In beamforming, a plane wave approximation is assumed, and the known apparent velocity and BAZ of the seismic event are used to align and sum the coherent parts of the seismic signals in every station to form a beam.

Seismic signal approaches the array as a plane wavefront. The signal reaches different sites at different times and these time delays τ_j has to be solved. Figure 2 illustrates an array of seismic stations and a plane wavefront propagating on a 2-dimensional surface.

The above time delays τ_j between the stations can be calculated by knowing the locations of the station $\boldsymbol{r_j}$ and the apparent slowness $\boldsymbol{s}$ of the seismic phase.

$$\tau_j = \boldsymbol{r_j} \cdot \boldsymbol{s} \tag{2.8}$$

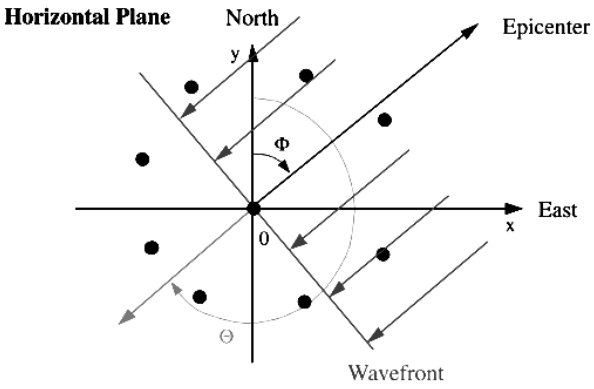


Figure 2: Illustration of a plane-wave approaching the array, the seismic phase will reach each of the stations at different times, dots denote the seismic stations. Θ is the direction to the epicenter, i.e., BAZ. Φ is the propagation angle of the wavefront. The original illustration from Schweitzer et al. (2012).

Simple beamforming formula is as follows:

$$b(t) = \frac{1}{M} \sum_{j=1}^{M} w_j(t + \tau_j) \tag{2.9}$$

where M the number of stations, w_j is the data sample of the station j at the time $(t + \tau_j)$, that is after time delay is applied.

Figure 3 visualizes the relation between the incidence angle and apparent velocity. The seismic phases approach the array in any incidence angle i between 0 to 90 degrees. The vertical and horizontal apparent velocities are functions of the incidence angle and actual propagation velocity v_c.

$$v_c = v_{app,h} \sin i = v_{app,z} \cos i \tag{2.10}$$

Figure 4 demonstrates the process of applying the delay times to form a beam.

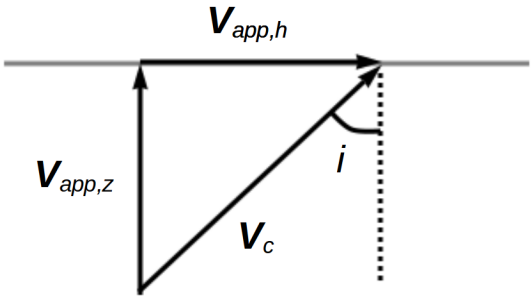


Figure 3: The wavefront approaches the station with incidence angle i and velocity v_c. $v_{app,h}$ and $v_{app,h}$ are apparent velocities.

Beamforming will increase the SNR. The increase of SNR called gain G, is related to the number of stations M in the array as follows:

$$G = \sqrt{M}, \tag{2.11}$$

where M is the number of stations. (Schweitzer et al., 2012) For FN network this means theoretical gain values of $\sqrt{6} \approx 2.45$ and $\sqrt{9} = 3$ for 6 and 9 stations, respectively.

In Figure 5 the suppression of the noise and increase of the SNR can be seen. Of course, seismic phases with diverging slowness values are also suppressed. Especially the arrival of the seismic phase around 5000 seconds mark has become visible. This naturally means that the mentioned phase has to have the same slowness vector as the triggered phase which is used to solve the slowness vector in the first place.

In this thesis, only surface waves are studied which means that the apparent velocity is actually the real effective velocity. To solve the apparent slowness and BAZ, beampacking method is used.

The deviations in elevations of the sensors are not taken into account. As stated in Table 1, the elevations of stations of the FN network range from 60 to 365 meters. The maximum difference of 305 meters being much smaller than the aperture of the FN network and the wavelenght range considered, the effects of elevation changes are

negligible. In addition, as only surface waves are of interest and no signals are expected to emerge from below of the array, the significance of the differences in elevations is further reduced. Also, the elevation deviations are neglected when calculating array transfer functions.

2.4.4 Beampacking

One of the methods to resolve the time delays for beamforming is beampacking. A beam is calculated using different slowness $\boldsymbol{s}$ values. For each slowness value (s_x, s_y) the delay is applied and a beam is formed. The power of the beam is then calculated. The result is relative power as a function of s_x and s_y. The maximum of this function should be the slowness of the seismic phase in question, which also includes the BAZ information. (Schweitzer et al., 2012)

More about the other methods to estimate the apparent slowness and BAZ of the event can be found in Schweitzer et al. (2012).

2.4.5 Signal and noise cross-correlation and gain

In order to analyze the coherency of the signals and noise inside the area of the array, signal and noise cross-correlation values of each unique pair of stations can be calculated. In the case of 9, stations this means 36 unique pairs. The correlation values are plotted as a function of interstation distance, i.e., the distance between two stations. As there are 36 station pairs, there are also 36 different distances.

We can calculate the SNR gain G as follows:

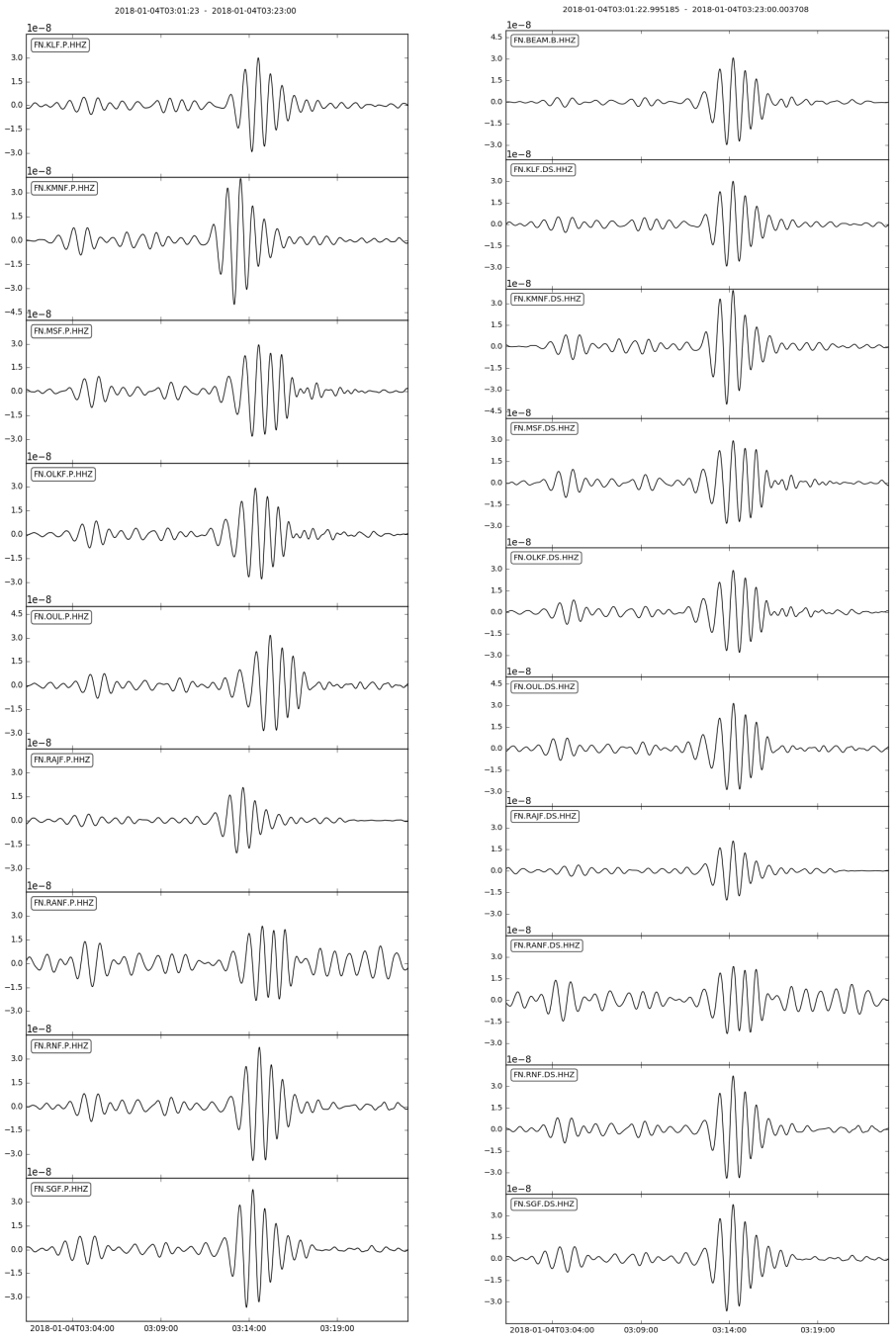
$$G^2 = \frac{\sum_{ij} C_{ij}}{\sum_{ij} \rho_{ij}} \tag{2.12}$$

where C_{ij} is the signal cross-correlation and ρ_{ij} is the noise cross-correlation. Index pairs ij correspond to the combinations of each of the two stations of the array. The FN network consists of nine stations so the amount of different station pairs is 81. For each of these pairs, we must calculate the signal and noise cross-correlations. (Schweitzer et al., 2012)

In theory, $C_{ij} = 1$ for all i and j and $\rho_{ij} = 1$ if $i = j$ and $\rho_{ij} = 0$ if $i \neq j$. Passing these values to Equation 2.12 will produce the theoretical gain formula (Equation 2.11).

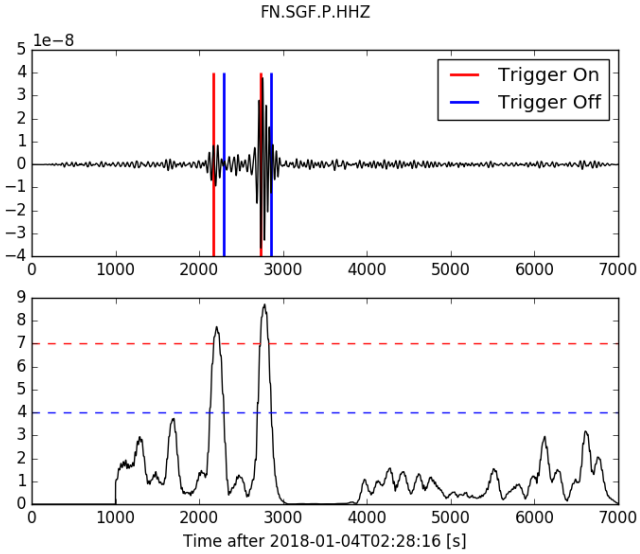
$$G^2 = \frac{\sum_{ij} C_{ij}}{\sum_{ij} \rho_{ij}} = \frac{M^2}{M}, G = \sqrt{M}$$

We will follow the practice described by Schweitzer et al. (2012). In signal and in noise calculations, the data should be filtered using appropriate bandpass filter and correlations should represent particular frequency ranges. The correlations will be plotted as a function of the distance between sensors i and j.

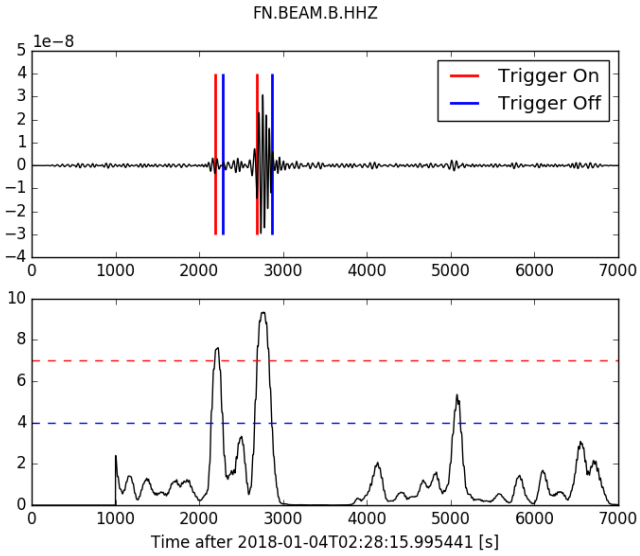


(a) (b)

Figure 4: (a) Seismic signals before time delay correction. (b) Seismic signals after time delays are taken into account. The seismic phases are aligned and the beam (the topmost signal) is calculated. The information of this seismic event (4.1.2018) in Table 5, 15mHz to 25mHz bandpass filter is used.



(a)



(b)

Figure 5: (a) Trigger station (SGF) STA/LTA trigger plot. (b) STA/LTA trigger plot of the beam. In both cases, the lower curve is the STA/LTA ratio. Red and blue dashed lines are the trigger's on and off threshold values. The information of this seismic event (4.1.2018) in Table 5, 15mHz to 25mHz bandpass filter is used.

3 FN network

In Tables 1 and 2 location coordinates of stations of the FN network are shown. Table 3 describes the instrumentation of each station. Every station is equipped with the 3-component broadband seismometer.

Figure 1 shows the locations of the stations in a map. For array methods, locations of nearby stations of other networks are shown. These stations could be used to form virtual arrays.

FN network is relatively large. If we consider it as an array, it has an aperture of around 455 kilometers.

Table 1: FN network station locations, all stations located inside Finland, latitude and longitude are in WGS84 coordinate system. Elevation seems not to be in GRS80 system. Station coordinates from International Seismological Centre (2015), 29.6.2018, URL: `http://www.isc.ac.uk/registries/`.

Code	Station name	Latitude (°)	Longitude (°)	Elevation (m)	Depth (m)
KLF	Kolari	67.23470	23.96400	209.2	1.0
OUL	Oulu	65.08530	25.89640	60.0	
SGF	Sodankylä	67.44211	26.52611	180.0	
RNF	Rovaniemi	66.60900	26.01350	171.0	
OLKF	Oulanka	66.32060	29.40030	315.0	5.0
MSF	Maaselkä	65.91131	29.04019	365.0	
RANF	Ranua	66.01470	26.78860	190.0	2.0
RAJF	Raja-Jooseppi	68.47512	28.31099	238.4	11.4
KMNF	Kaamanen	69.14895	26.99611	246.0	21.7

3.1 Noise characteristics

In Figure 6 shows the vertical component PPSDs of each station in the FN network. The mean of each PPSD is plotted in Figure 7 to ease the comparison between stations.

Table 2: FN network, coordinates presented in Table 1 transformed to ETRS-TM35FIN (datum ETRS89) coordinate system using Coordinate Transformation Service provided by Finnish Geodetic Institute (2007), 20.7.2018, URL: `http://coordtrans.fgi.fi/transform.jsp`.

Code	Northing (m)	Easting (m)
KLF	7460746.9021	368943.2801
OUL	7218414.2619	448125.9896
SGF	7480744.1433	479713.2580
RNF	7388138.9186	456297.8850
OLKF	7357709.3243	607550.7153
MSF	7311534.7213	592905.5918
RANF	7321564.7123	490410.8173
RAJF	7596408.7737	553676.5587
KMNF	7670970.8868	499845.4757

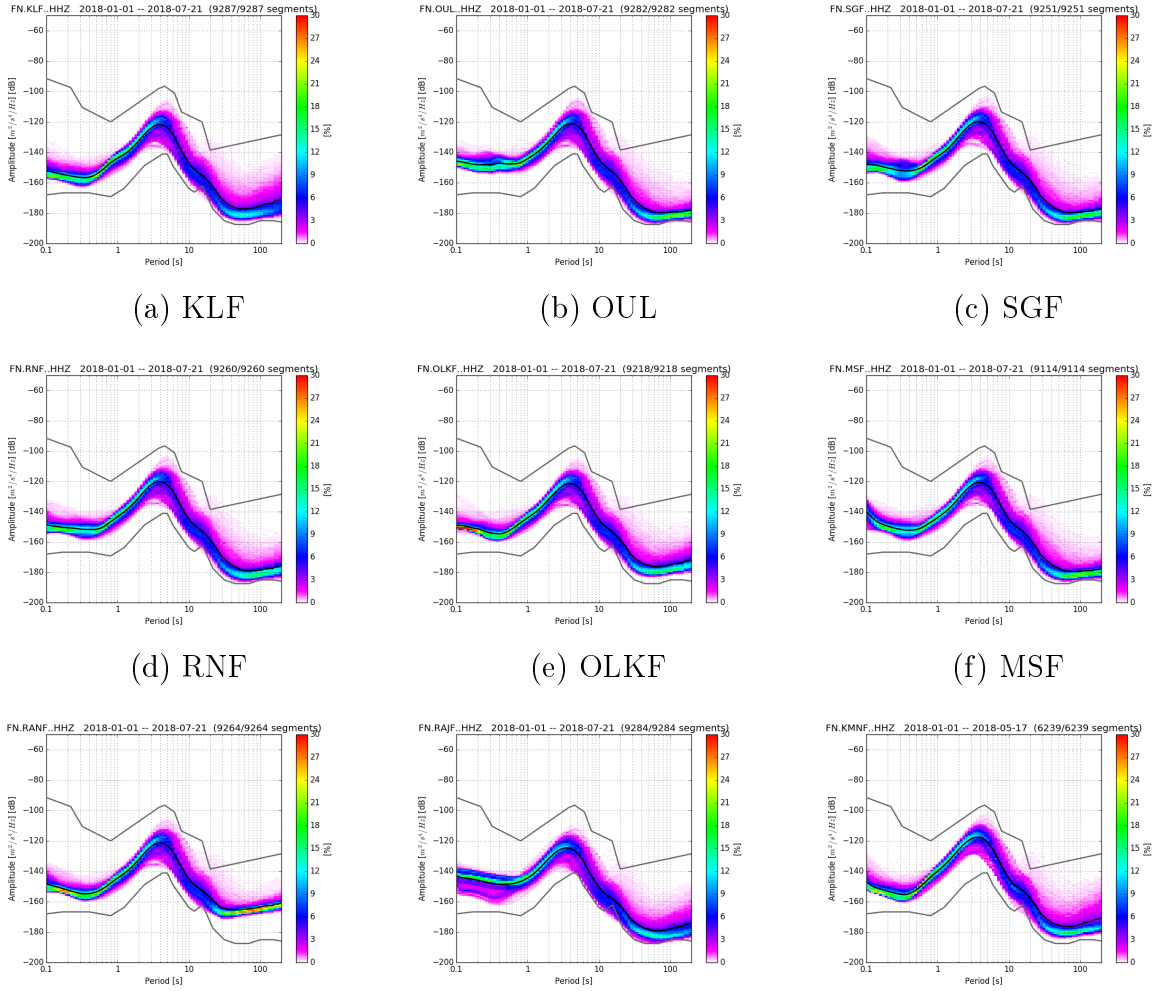
Table 3: The instrumentation of the FN network.

Code	Sensor	Digitizer
KLF	Nanometrics Trillium 120PA	Earth Data PS6-24
OUL	Streckeisen STS-2	Earth Data PS6-24
SGF	Streckeisen STS-2	Earth Data PS6-24
RNF	Streckeisen STS-2	Earth Data PS6-24
OLKF	Nanometrics Trillium Posthole	Earth Data PS6-24
MSF	Streckeisen STS-2	Earth Data PS6-24
RANF	Nanometrics Trillium 120PA	Nanometrics Centaur 3ch
RAJF	Nanometrics Trillium Posthole	Nanometrics Centaur 3ch
KMNF	Nanometrics Trillium Posthole	Nanometrics Centaur 6ch

To calculate the PPSDs, continuous seismic data from dates 1.1.2018 to 21.7.2018 was used.

From these plots, we can see that overall noise levels in the FN network are quite low. As for array applications in this study, we are interested in noise levels at periods over 40 seconds as the FN network as a relatively large array is most suitable for detecting

longer period seismic phases. In the particular period range between 40 seconds to around 200 seconds is very close to NLNM.



(a) KLF (b) OUL (c) SGF

(d) RNF (e) OLKF (f) MSF

(g) RANF (h) RAJF (i) KMNF

Figure 6: PPSD plots of vertical channels of FN network stations, black line denotes mean values.

3.2 Array transfer function

In order to compare the transfer function of FN network to Yellowknife transfer function represented by Schweitzer et al. (2012), using Obspy package by The ObsPy Develop-

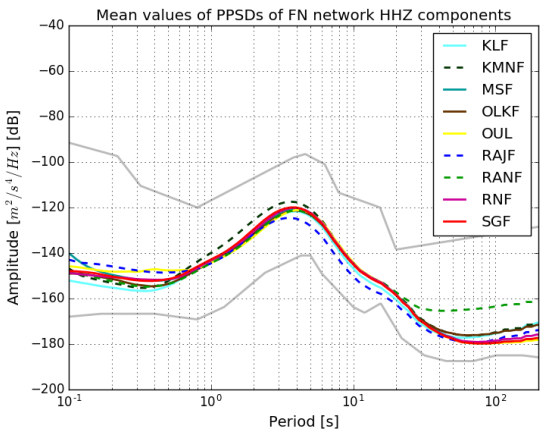


Figure 7: Mean values of PPSD plots of stations of the FN network, vertical components.

ment Team (2017), locations have to be presented in meters and multiplied by 2π. The intensity is then converted into decibels by applying base 10 logarithm and multiplying by 10. The presentation and the units of the transfer functions are equivalent to the ones used by Wang (2002).

In Figure 8 the array transfer functions of 6-station and 9-station FN network and some virtual arrays with different station configurations are presented. The additional stations for the virtual arrays on top of FN stations are selected from the ones in Figure 1. The deviations in elevations of stations are not taken into account when calculating the transfer functions.

The transfer function is calculated with every combination of 1-5 external stations and the transfer functions producing the lowest sidelobe maxima are presented in Figure 8. In Figure 8, the transfer functions are produced using, in a sense, optimal selections of external stations. With 5 external stations, the inclusion of the HE.OUF stations produced the lowest maxima, but since the inclusion of this station also increases the array aperture, the calculation was made again excluding the HE.OUF station.

3.3 Array resolution and wavenumber passband

The resolution of the array is determined by the aperture and the distribution of the stations. When using a plane-wave approximation, the highest error in the propagation direction is achieved when the wave is arriving perpendicularly to the largest dimension of the array. (Johnson and Dudgeon, 1993)

Next, the resolution considerations presented by Wang (2002) are followed and applied to the FN network. The width of the mainlobe defines the minimum resolvable wavenumber k_{min}. The mainlobe width itself is defined by the -3 dB contour line. If the mainlobe is not circular the width is the average of the lengths of the semi-minor and semi-major axes. The maximum resolvable wavenumber k_{max}, on the other hand, is defined by the half distance from the mainlobe center to the center of the closest grating lobe.

In the case of the FN network with 9 and 14 stations (cases (b) and (f) in Figure 8)

$$k_{min,9} = 9 \cdot 10^{-3}\ 2\pi/km,\ k_{max,9} = 40 \cdot 10^{-3}\ 2\pi/km$$

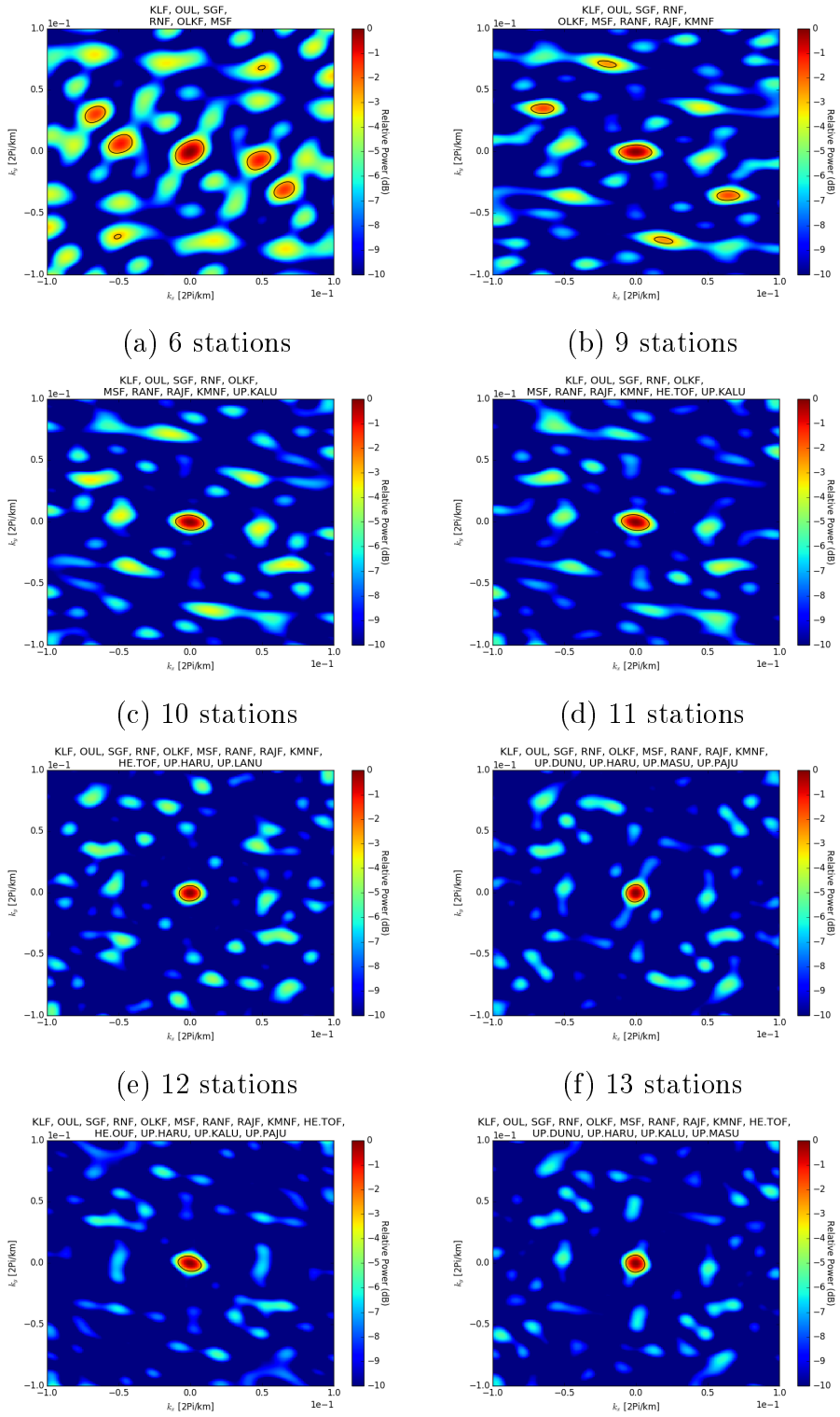
$$k_{min,14} = 7 \cdot 10^{-3}\ 2\pi/km,\ k_{max,14} = 60 \cdot 10^{-3}\ 2\pi/km.$$

The above approximate values are taken from transfer functions (b) and (f) in Figure 8.

The wavenumber resolution values define a wavenumber passband which describes frequency-slowness values best detected by the array. The values are plotted into logarithmic frequency-slowness plot using the relation

$$k = 2\pi/\lambda = 2\pi f s. \tag{3.1}$$

Figure 9 illustrates wavenumber passbands corresponding 9 (a) and 14 (b) station configurations respectively. The additional vertical lines denote the frequency values of interest. The passband describes the slowness and frequency values of the seismic phase that are most detectable by the array.



(a) 6 stations

(b) 9 stations

(c) 10 stations

(d) 11 stations

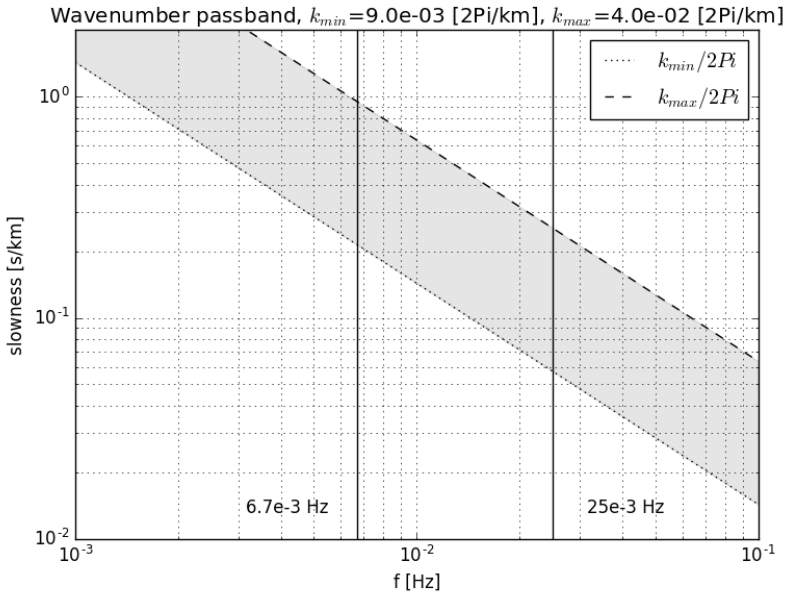
(e) 12 stations

(f) 13 stations

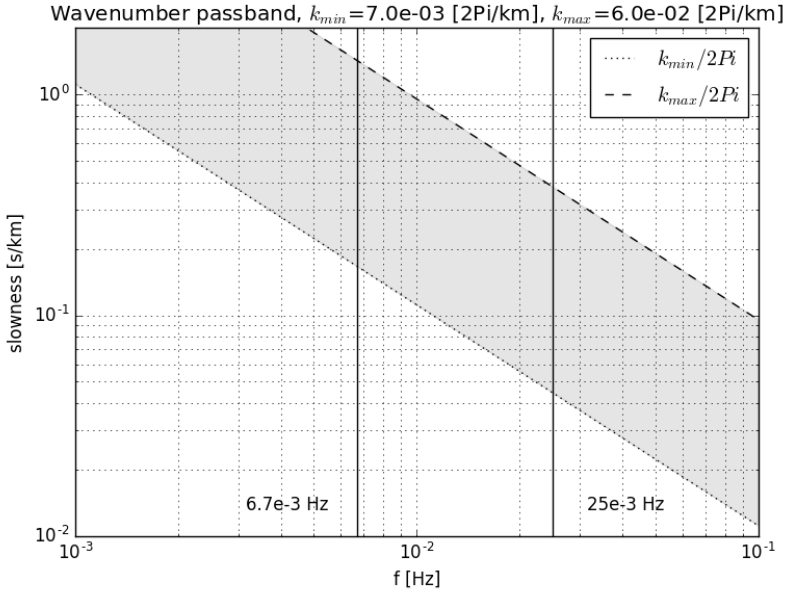
(g) 14 stations

(h) 14 stations

Figure 8: The FN network array transfer functions achieved using different configurations, the black lines are the -3 dB contour lines.



(a) 9 stations



(b) 14 stations

Figure 9: The wavenumber passband of the FN network with 9 initial stations (a). The wavenumber passband of the virtual array of FN stations and 5 external stations (b). The five external stations are the same as in Figure 8 (h).

4 Collected data

4.1 Earthquake location tests

In this chapter all coordinates (latitude and longitude of the earthquake epicenters) are in WGS84 coordinate system.

To test the earthquake locating capabilities of the FN array, earthquake catalogs using event information from 1.1.2018 to 21.7.2018 were created. The event information was extracted from Orfeus Data Center (2018) for the global catalog and International Seismological Centre (2018) for the arctic catalog. For the arctic catalog, only events originating from the area of longitude from 29° W to 74° E and latitude from 43° N to 85° N are included, though only one event with latitude lesser than 60° N exists in the catalog. The localization tests are presented in Chapters 4.1.1 and 4.1.2.

From the catalogs, events too close to each other in time are excluded. Time timespan in which the multiple events are not allowed is exclusion timespan in Table 4. The actual timespan where one event should be found is timespan in Table 4. Both timespan and exclusion timespan start at the origin time of the event described in the catalogs. The timespans are defined so that the seismic phase of interest, in this case, surface waves with the velocity of around 4 km/s, is recorded from events originating from anywhere in the area of interest.

From the continuous seismic data, an interval of data which contains the event is extracted. Before and after the actual timespan, an offset (start and end offsets in Table 4) is added. These offsets are necessary for the signal processing for example for STA/LTA trigger, tapering, and filtering. The data interval consists of vertical component data of every available station.

The extracted data interval is cosine tapered, the length of the taper in Table 4. After tapering, a bandpass filter of 15mHz to 25mHz is applied.

STA/LTA trigger is used to detect the seismic phase. The trigger is only applied to the trace of the trigger station Table 4. In the case of global catalog, SGF station was used as a trigger station, since it is a central station and equally remote to the teleseismic sources. Parameter values for the trigger are described in Table 4. If the 'on' trigger lies inside the timespan, the phase is found. If multiple triggers are found, the latest one is chosen.

After triggering, the time interval corresponding to the seismic phase is extracted. The newly extracted slice starts 260 seconds before the 'on' trigger and ends 130 seconds after 'off' trigger. These offsets are chosen to include the whole waveform corresponding to the extracted phase in all stations.

Only vertical components were used, and only the surface wave phases were expected to be recorded, thus the actually recorded seismic surface waves were probably Rayleigh waves.

The data which contains the filtered vertical component data of each available station and only the seismic phase in question, proceeds to the beampacking.

An uniform slowness grid of 10201 points and slowness values $s_x, s_y \in [-0.6; 0.6]$ s/km is created for beampacking. In each grid point, using beamforming (Equation 2.9) with slowness values s_x, s_y and calculating the power of the corresponding beam will give the value of point s_x, s_y in the resulting beampacking plot.

The central station of the array (SGF in all cases) will always have the time delay of zero and resulting BAZ will be calculated with respect to the location of the center station.

The power value in point s_x, s_y is the mean value of power spectral density (PSD) of the beam between frequencies from 15mHz to 25mHz. The PSD is calculated without overlap and the length of the fast Fourier transform block in samples is the largest power of two which is still lesser or equal than the number of samples in one trace divided by 4.

The location of the maximum in the beampacking result plot will give the estimate of the slowness of the phase. The BAZ of the event can be calculated from the location of the maximum as well.

The newly resolved BAZ values are compared to the actual BAZ values calculated using the earthquake catalogs. From now on, the term calculated BAZ will be used for BAZ values calculated using beampacking and the term original BAZ will stand for BAZ values calculated using earthquake catalogs.

Table 4: Parameter values used in earthquake location tests in case of different catalogs.

Parameter	Global catalog	Arctic catalog	North Korea
exclusion timespan (s)	5000	1200	5000
timespan (s)	5000	1200	5000
start offset (s)	1000	1500	1000
end offset (s)	1000	1500	1000
taper length (s)	500	120	500
STA/LTA, STA length (s)	100	100	100
STA/LTA, LTA length (s)	1000	1000	1000
STA/LTA, 'on' threshold	7	5	7
STA/LTA, 'off' threshold	4	1	4
trigger station	SGF	RAJF	SGF
center station	SGF	SGF	SGF

In Figure 10, the results of earthquake locating tests using global catalog are visualized. Cumulative distribution function of the results of global catalog is plotted in Figure 11. The horizontal axis in result and CDF plots is the absolute value of BAZ deviation

between calculated BAZ and original BAZ. Magnitude, depth, distance and original BAZ of the event together with the apparent velocity of the seismic phase are plotted with respect to above-mentioned deviation. In Figure 10, each event is plotted in each subplot, creating 5 data points per event in total, all with the same BAZ deviation value.

4.1.1 Global catalog

The event information was retrieved from Orfeus Data Center (2018), data downloaded 17.10.2018, URL: `https://orfeus-eu.org/webdc3/`. The data set consists of days from 1.1.2018 to 21.7.2018 and includes 2647 events located all around the world. 2066 events are excluded due to the limit of only one event per timespan. The trigger did not detect 414 of the remaining events, leaving 167 events which proceed to the localization tests, in which the BAZs of the events are calculated.

Figure 10 presents the results of earthquake locating tests using the global catalog. It is visible that as the magnitude increases the accuracy of localization increases, as the higher magnitude events are more focused on smaller BAZ deviation values. And in fact, the BAZ deviation of events with magnitude greater than 5.5 was at most 12°. Also, little indication of a decrease in localization accuracy can be seen as the distance to the epicenter increases.

In velocity plot, almost all phases had the apparent velocity of around 4 km/s, but as it may be seen there were registered events with higher velocities also. In these cases, the extracted phase probably was not the surface wave phase. These phases might be body waves approaching the array with high apparent velocities or just artefacts from the signal processing.

Figure 11 is the cumulative distribution function of the deviation values. From this figure it can be seen that more than 75 percent of the events have the deviation of fewer

than 10 degrees.

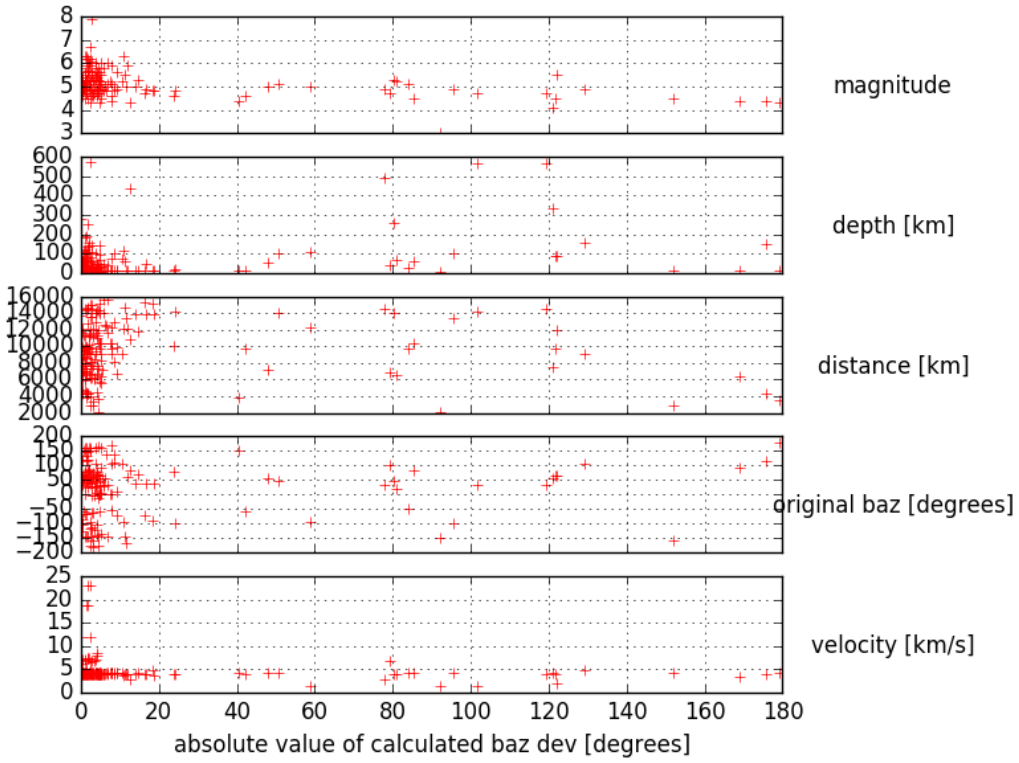


Figure 10: Localization results using the global catalog, 167 events. All parameters except the BAZ deviation are from the earthquake catalog.

Figure 12 shows the trigger status and distance and magnitude relation of 581 events of global catalog. In case of these events, the general lack of low magnitude events at long distances, is probably due to exclusion of events too close each other in time. It may be observed that events from smaller epicentral distances and with higher magnitude are more likely to be successfully triggered. In the case of the global catalog, every event with a magnitude over 6.2 (5 events) was successfully triggered. The catalog contains only events with a magnitude over 3, and one of such events was also successfully triggered.

Example event

Table 5 contains the information about the global catalog example event.

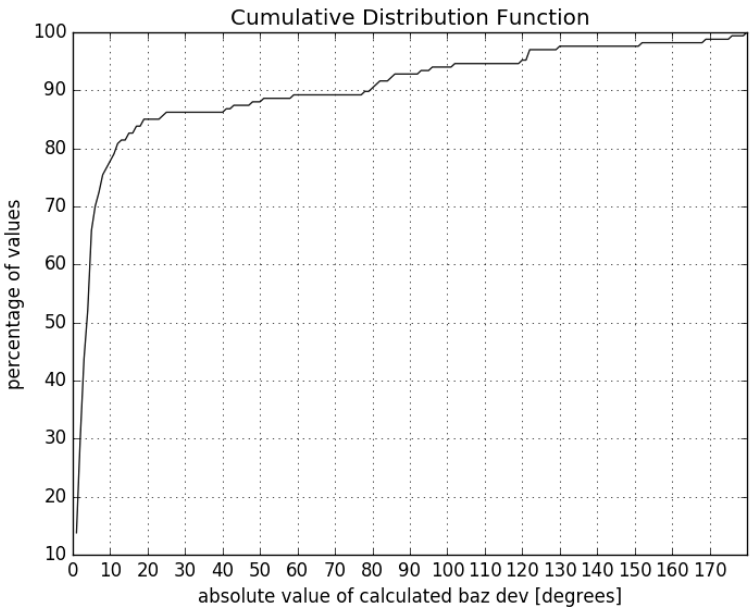


Figure 11: Cumulative distribution function of the results of the BAZ evaluation for the events from the global catalog, 167 events.

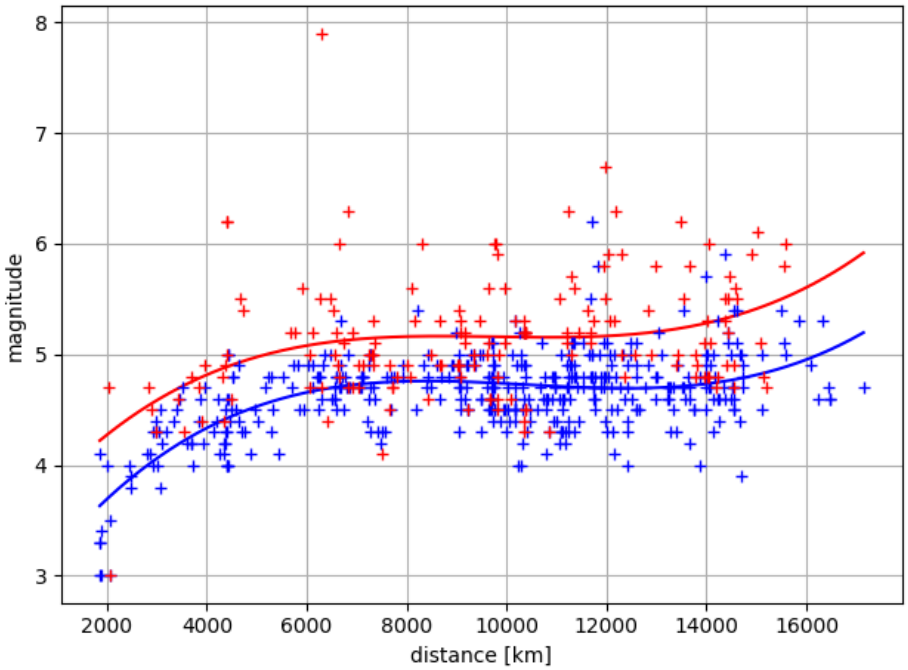


Figure 12: Distance and magnitude relation of successful (167) and failed (414) events of global catalog. Red datapoints are successfully triggered events.

Table 5: Example event of the global catalog.

origin time (UTC)	2018-01-04 02:44:56
magnitude	5.0
depth (km)	47.0
region	Near East Coast of Kamchatka
latitude (°)	53.26
longitude (°)	159.77
distance (km)	6064
original BAZ (°)	32.4

Figure 13 shows the result of beampacking analysis for the example event of global catalog. Figure 4 demonstrates the beamforming of this particular example. In Figure 4 (a) seismic signals are not aligned. After resolving the apparent slowness and BAZ using beampacking (see Figure 13) and applying the appropriate delay times, coherent parts of the signals are summed to form a beam. The beam is the topmost signal in the plot (b) in Figure 4, other seismographs in plot (b) are the signals after applying the delay times.

The original BAZ of the event is 32.4° (see Table 5) and the calculated BAZ is 32.91° (see Figure 13), which results in the deviation of BAZ of around 0.51°.

4.1.2 Arctic catalog

The event information was extracted from ISC Bulletin (International Seismological Centre, 2018), data downloaded 22.10.2018, URL: `http://www.isc.ac.uk/iscbulletin/search/catalogue`. The data set consists of days from 1.1.2018 to 21.7.2018. Events are located inside the area of longitude from 29° W to 74° E and latitude from 43° N to 85° N. The catalog contains only one event with latitude lesser than 60° N and after triggering it is excluded. The arctic region is considered to start at latitude over 60° N.

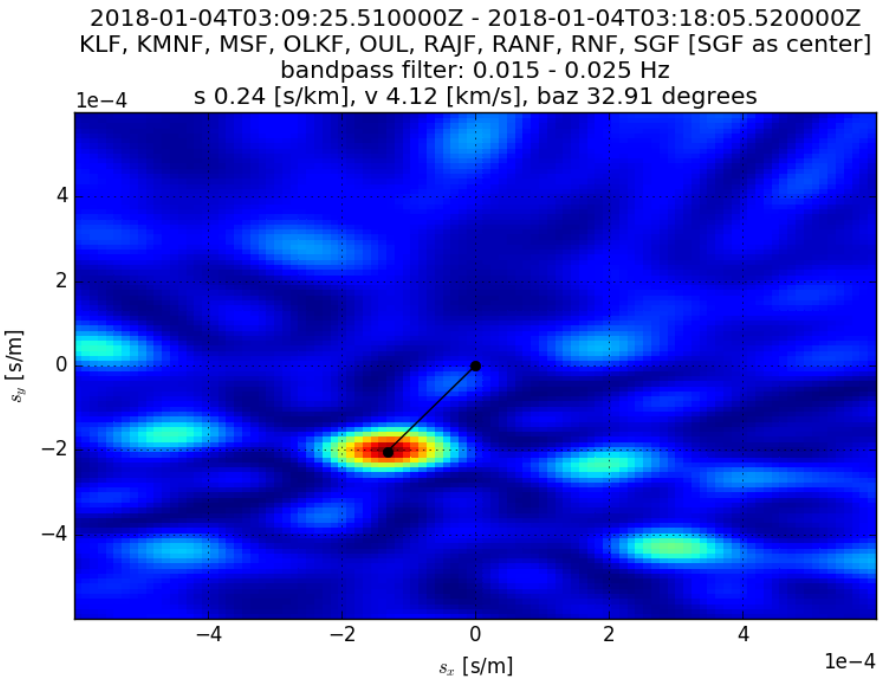


Figure 13: The results of beampacking of the example event of the global catalog (Table 5), a relative power of the beam as a function of lowness vector (s_x, s_y). the black line denotes the maximum, the resulting BAZ is 32.91°, 15mHz to 25mHz bandpass filter is used. The values except the BAZ deviation are from the earthquake catalog.

As the FN network is located relatively north, locating arctic events, such as tectonic events originating from the Arctic or glacial events from the Greenland ice sheet, might be of interest. A similar event selection criteria, as in the case of the global catalog, were used. The same restrictions apply with changed parameters, see Table 4.

At the start, there were 690 events of which 125 are excluded for being too near each other in time. Trigger fails in 480 of the cases, leaving 85 events for localization tests. Four events had a magnitude of greater than 4.7 and all of these were succesfully triggered. The lowest magnitude of a succesfully triggered event was 0.6. Generally, the magnitudes of the events range from 0.3 to 5.4. The events in the arctic catalog were quite shallow and mostly originated from north and west.

A set of 85 events was used in localization tests, but considering the plane-wave approx-

imation, it is not very meaningful to analyze events located too close to the array. The localization of events with epicentral distances smaller than 1000 kilometers from the array center turned out to be problematic. From the 85 events, only 43 have epicentral distances over 1000 kilometers from the array center, the largest distance still being only around 2700 kilometers. In the case of the arctic catalog, events are generally too close to the array for the plane-wave approximation to work. After applying all the conditions and requirements, too few events from the arctic catalog were available for comprehensive analysis.

Compared to the global catalog, it can be noted that generally lower magnitudes make the localization more difficult. Even after adjusting the trigger to detect smaller events, the number of failed triggers is quite high.

Example event

Figure 14 shows the result of beampacking of the event. The figure describes the calculated BAZ of the event (-54.78°). As a result, original BAZ being -58.9° (see Table 6), the BAZ deviation is 4.12°.

Table 6: Example event of the arctic catalog.

origin time (UTC)	2018-04-25 00:05:01
magnitude	3.4
depth (km)	0.0
region	Near Jan Mayen in the Norwegian Sea
latitude (°)	70.83
longitude (°)	-6.87
distance (km)	1361
original BAZ (°)	-58.9

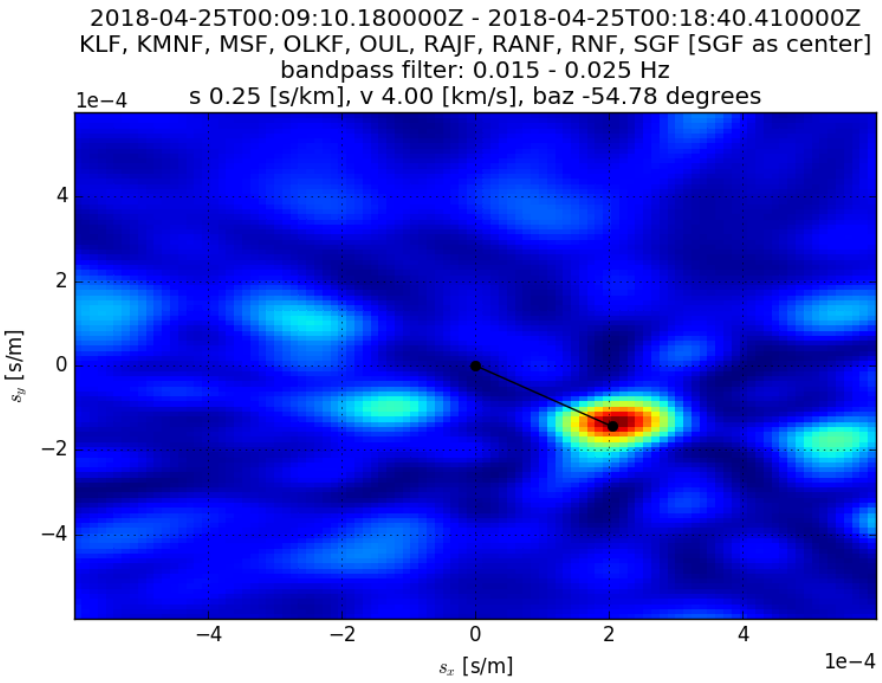


Figure 14: The result of beampacking of the example event of the arctic catalog (Table 6), a relative power of the beam as a function of slowness vector (s_x, s_y), the black line denotes the maximum, the resulting BAZ is -54.78°, 15mHz to 25mHz bandpass filter is used.

4.1.3 North Korea nuclear test 3.9.2017

On 3 September 2017 North Korea conducted a nuclear test (NORSAR, 2017). As the detection and localization of nuclear detonations were one of the primary reasons behind the development of seismic arrays and array methods, the detection of such events using the FN network, is of interest.

The localization was done similarly as in the case of the global catalog, see the used parameter values in Table 4. The event information was similar and it was downloaded from Orfeus Data Center (2018), 10.12.2018, URL: `https://orfeus-eu.org/webdc3/`.

Figure 15 shows the results of beampacking of the event. The deviation in BAZ is 2.23°, see original BAZ and calculated BAZ in Figure 15 and in Table 7.

Table 7: North Korea nuclear test 3.9.2017.

origin time (UTC)	2017-09-03 03:30:01
magnitude	6.3
depth (km)	1.0
region	Northern North Korea
latitude (°)	41.30
longitude (°)	-6.87
distance (km)	6336
original BAZ (°)	61.2

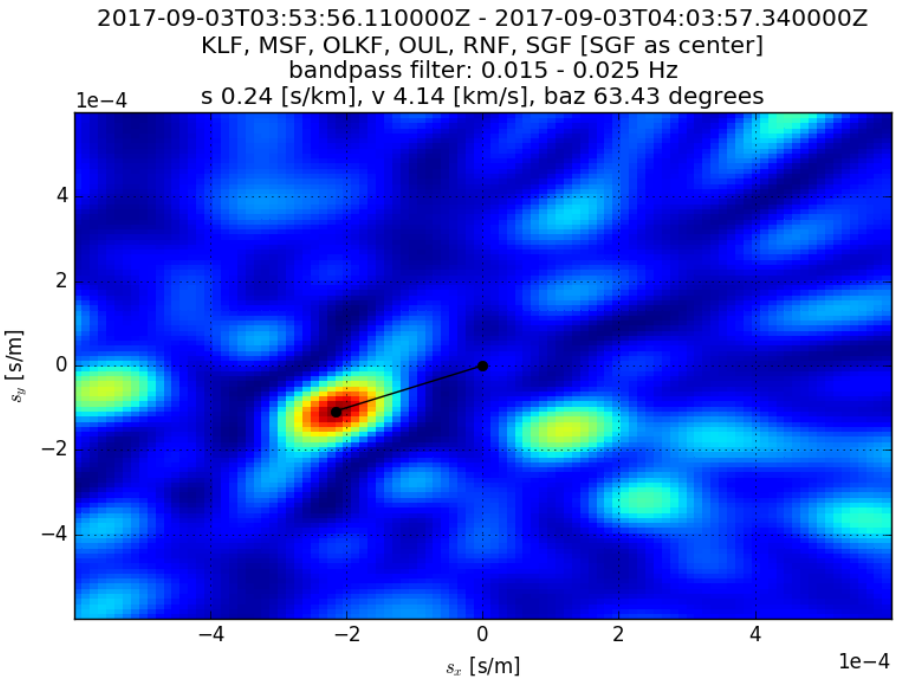


Figure 15: The result of beampacking of the North Korea nuclear test event (Table 7), a relative power of the beam as a function of slowness vector (s_x, s_y), the black line denotes the maximum, the resulting BAZ is 63.32°, 15mHz to 25mHz bandpass filter is used.

4.2 Signal and noise cross-correlation and achieved gain

In Figure 16 signal and noise cross-correlation values as a function of interstation distance are plotted. The waveform data is from the 167 events from the global catalog

(see page 31). The signal cross-correlation value is the median of these 167 events.

For noise cross-correlation values, continues seismic data between days 10.1.2018 and 21.7.2018 is divided into 15-minute-long blocks. The time interval must not contain seismic events. The events are located using the continuous data of SGF station and STA/LTA trigger (LTA length 500 seconds, STA length 50 seconds, 'on' threshold is 3 and 'off' threshold is 1). After triggering, 15120 blocks of seismic noise are acquired. The noise cross-correlation value is then the median of cross-correlation values of these blocks. There are 36 signal and noise cross-correlation values since there are 36 unique pairs of stations in 9-station array. Both signals and noise blocks are bandpass filtered using 15mHz to 25mHz frequency band. Figure 16 shows a quite good correlation of signals across the array. In the selected frequency band, the shape of the noise cross-correlation curve behaves similarly as presented in Schweitzer et al. (2012).

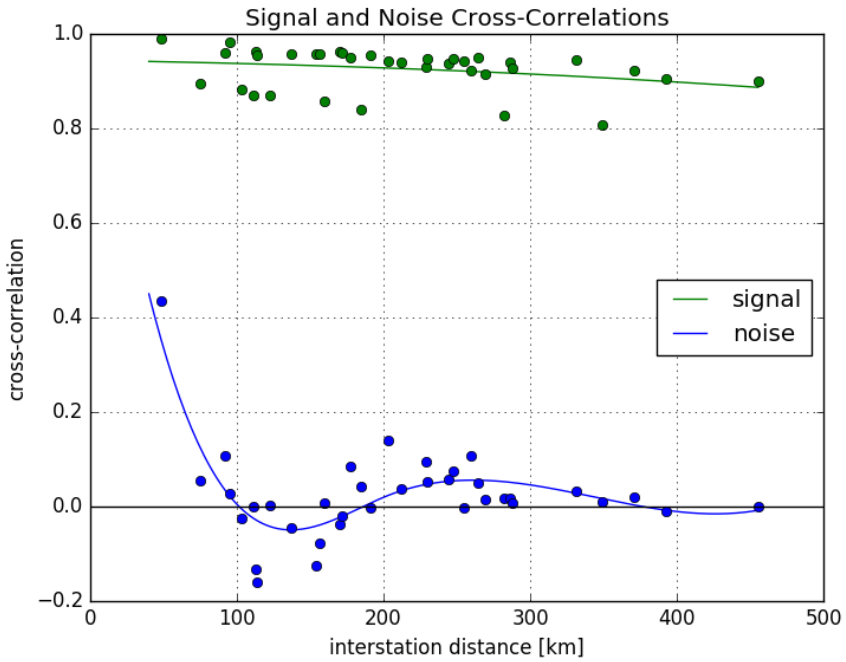


Figure 16: Signal (the green dots and interpolation) and noise (the blue dots and interpolation) cross-correlation, created using 36 station pairs, 167 seismic events, and 15120 15-minute long seismic noise blocks, 15mHz to 25mHz bandpass filter is used.

Using Equation 2.12 and the calculated signal and noise cross-correlation values (see Figure 16), the gain G of 2.66 is achieved. The theoretical value for maximum gain is 3 for an array of nine stations, as stated before in Chapter 2.4.3, page 15. As the gain is greater than 1, increase of SNR was achieved. The achieved gain is bound to the used frequency band.

4.3 Using array techniques to select locations for new seismic stations

In this chapter some considerations are presented to possibly take into account when selecting sites for new seismic stations in the FN network. New site locations are selected so that the inclusion of the station would reduce sidelobes in the array transfer function the most. This will expand the detectable wavenumber space of the array and thus improve for example the detection of seismic events with different back azimuths.

In this work, a regularly spaced grid representing the possible locations of a new station is defined in the study area. In this case, the grid consists of 119601 points, $x \in [240000, 650000]$ m and $y \in [7200000, 7780000]$ m. The grid point interval in the y-direction is 2 km and in x-direction 1 km. The grid is in ETRS-TM35FIN coordinates, and from now on all coordinates are in this coordinate system.

For each grid point, a new array transfer function is calculated assuming that the new station is located in this point. The value in this point will be the sidelobe maximum, which in practice is calculated by removing the mainlobe and taking the maximum of the remaining function. The transfer functions are calculated as those presented in Figure 8, that is $k_x, k_y \in [-0.1; 0.1]$ $2\pi/km$, respectively.

The minimum in this minimized sidelobe plot will be the optimal location for a new station. The new location has to be inside Finland and it must not increase the aperture

of the array. No other restrictions are taken into account, which means that the newly calculated location for the seismic station might not be suitable for deployment in other respects.

Next, calculations and minimized sidelobe plots for locations for new seismic stations are presented. Cases A, B, and C are different starting configurations of stations. In each case, two iterations for the most suitable location are calculated. In the second iteration the optimal location obtained after the first iteration is taken into account as a new station location.

The case A was calculated without any external stations, only the 9 stations of the FN network. Cases B and C were calculated using external stations of HE network. The stations in the case B were selected so that the aperture of the array would not increase.

Table 8 presents the optimal locations for new seismic stations after each iteration.

Figures 17 and 18 are resulting plots in the case A, first and second iterations and Figures 19, 20, 21 and 22 are resulting plots in cases B and C, first and second iterations, respectively.

In these plots, the values of global minimum and maximum are indicated. The analysis of these values shows that if the amount of stations increases and if the locations of the stations are not exactly aligned, the sidelobe maxima are generally reduced.

Without the previously defined restrictions, station A2 would be at the minimum south-east of OUL station and the location of C1 would be around Karasjok in Norway if locations outside Finland are considered as potential sites. Even though these two stations are not located in actual minima, it can be clearly seen that the sites after relocation are almost equal, i.e. the value of the new minimum is very close to the actual minimum.

In addition to finding new locations for stations, it is discovered that these kinds of plots will visualize the geometry of the current array configuration. For example, in

Figure 17 the alignment of the stations becomes more visible. There are three visible directions of alignment in Figure 17. In the northwest-southeast direction (the dark red parallel lines), all of the stations are aligned in certain intervals, and in the other two directions, every station except MSF and maybe SGF are aligned. The alignment, in this case, means that the stations are located in parallel lines with a certain spacing.

Table 8: The optimal locations for new stations based on minimizing the sidelobes, Coordinates in ETRS-TM35FIN system.

New station	Northing (m)	Easting (m)
A1	7434000	509000
A2	7506000	437000
B1	7294000	430000
B2	7348000	567000
C1	7500000	426000
C2	7536000	571000

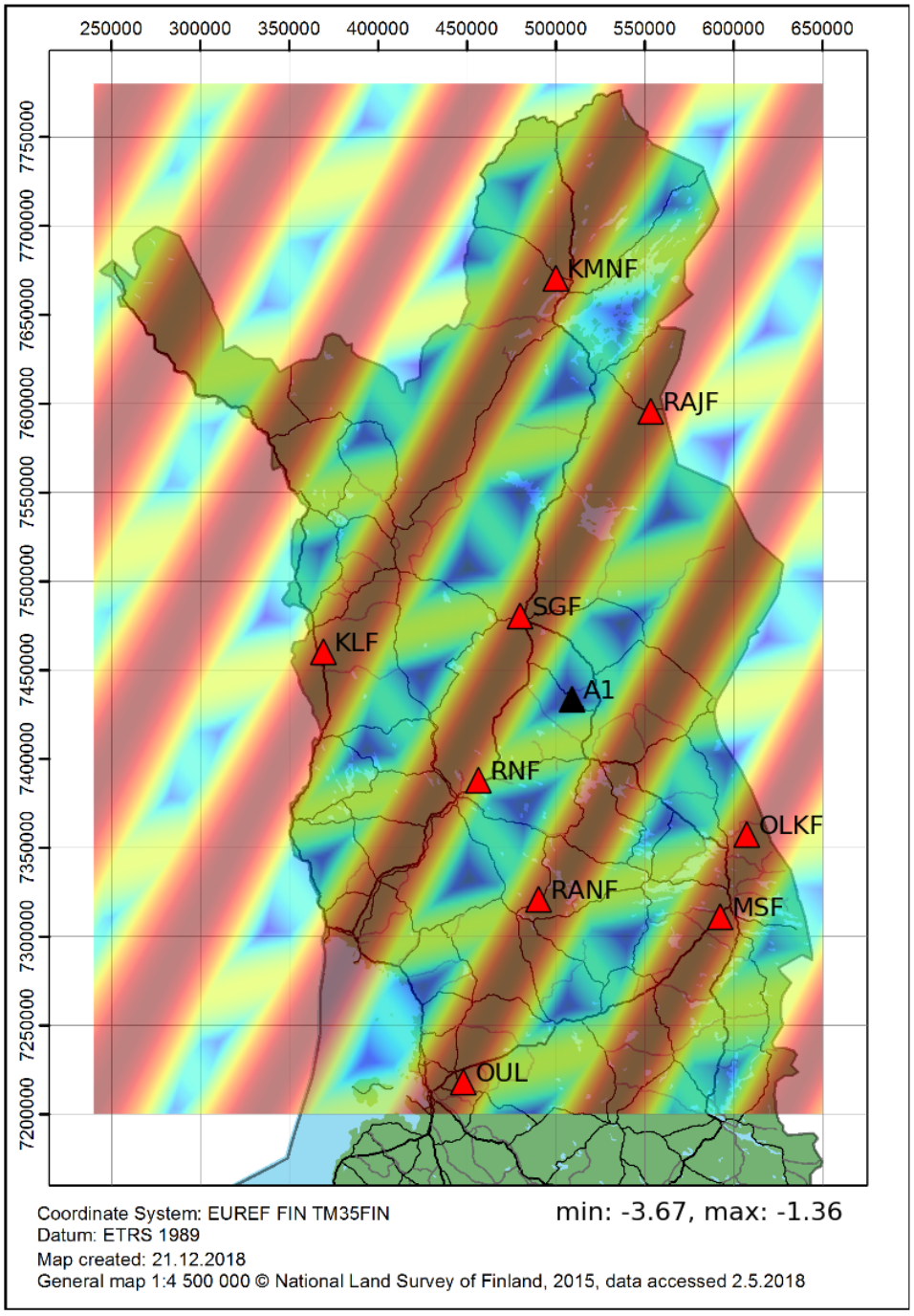


Figure 17: Minimizing sidelobes plot, the case A with only FN stations, the first iteration. The red triangles are the FN stations, and the black triangle is the optimal location for the new station. The values (relative power) range from -3.67 (the darkest blue) to -1.36 (the darkest red).

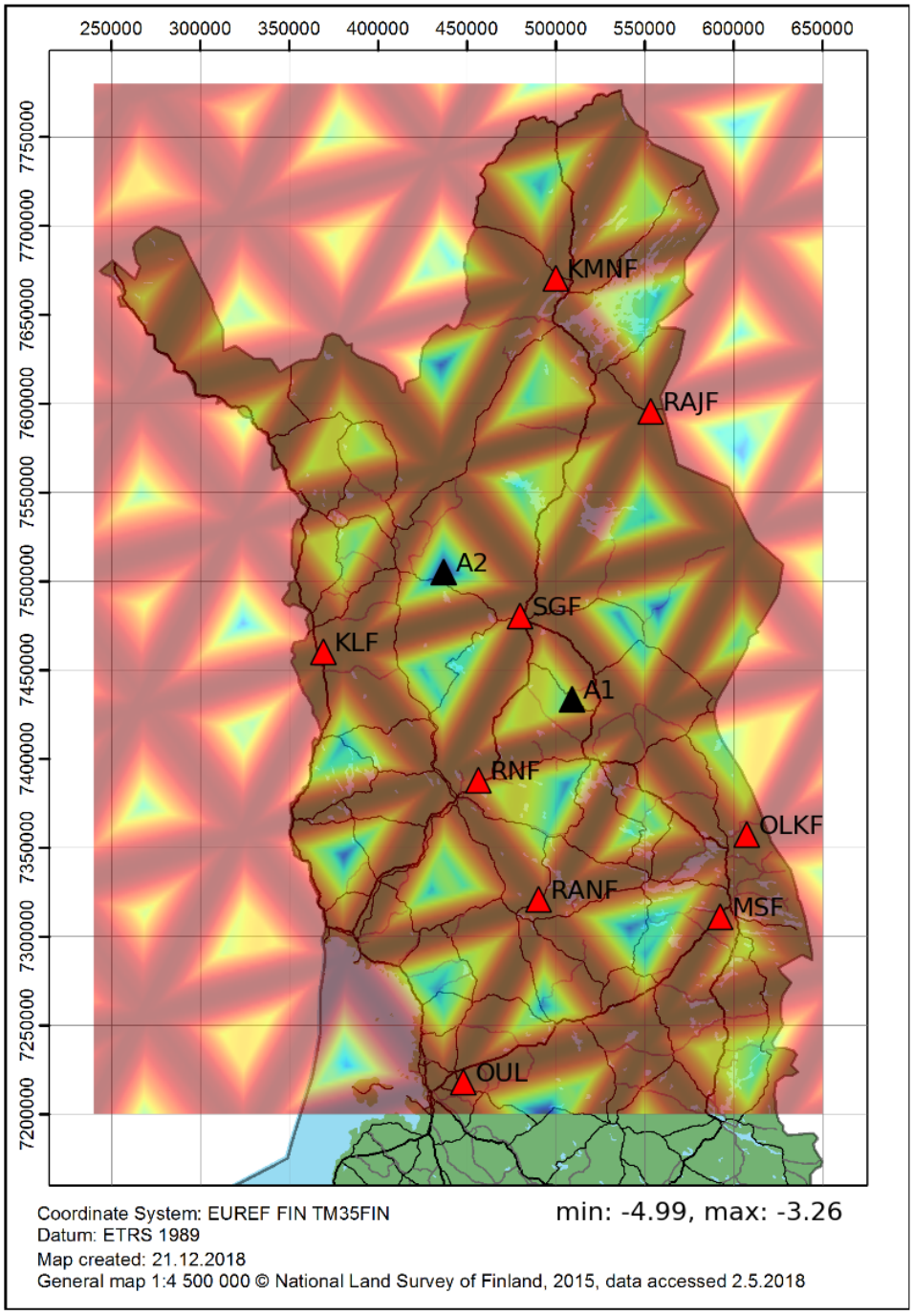


Figure 18: Minimizing sidelobes plot, the case A with only FN stations, the second iteration. The red triangles are the FN stations, and the black triangles are the optimal locations for the new stations (the first and second iteration). The values (relative power) range from -4.99 (the darkest blue) to -3.26 (the darkest red).

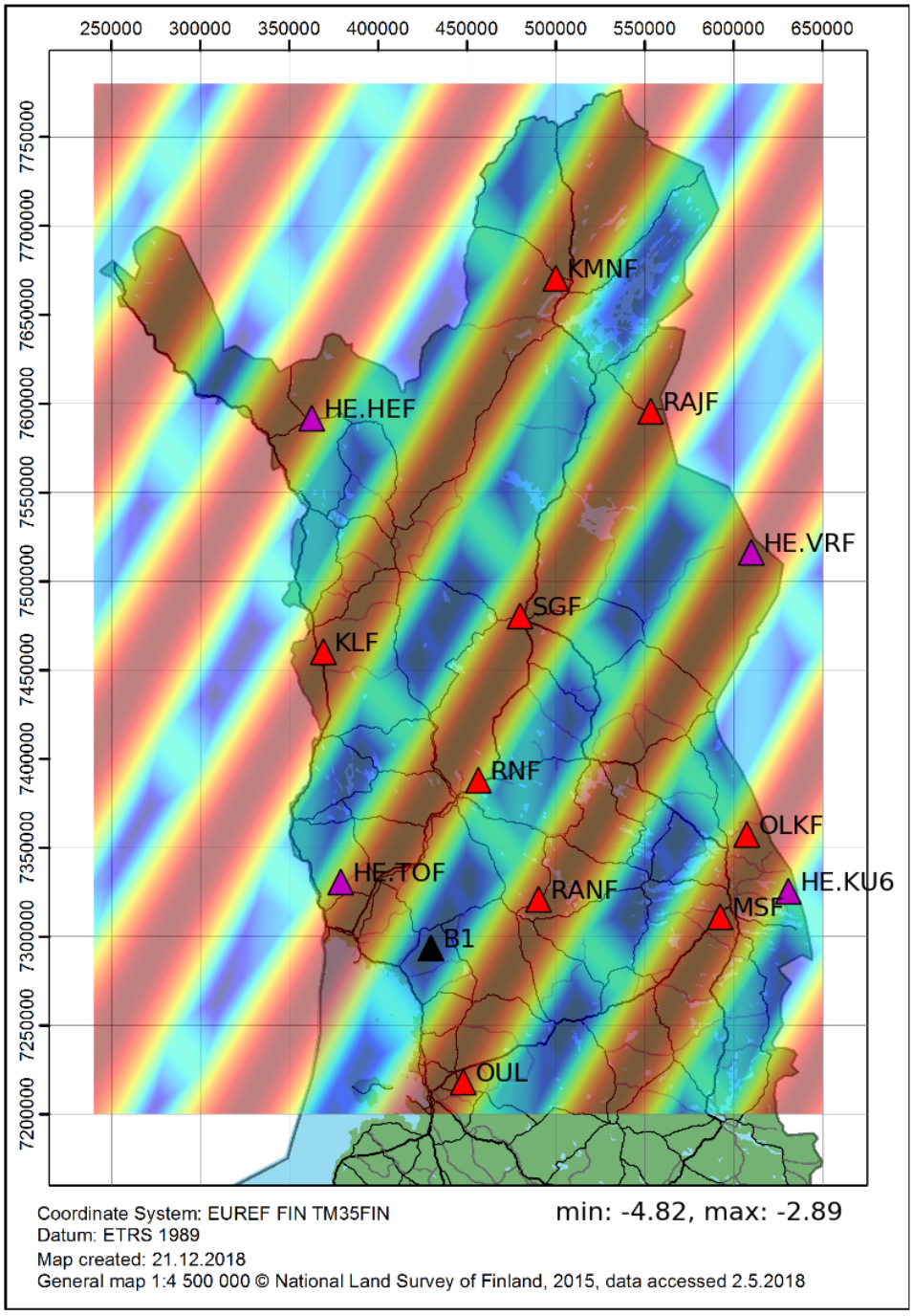


Figure 19: Minimizing sidelobes plot, the case B with 4 external stations, the first iteration. The red triangles are the FN stations, the magenta triangles are the used external stations, and the black triangle is the optimal location for the new station. The values (relative power) range from -4.82 (the darkest blue) to -2.89 (the darkest red).

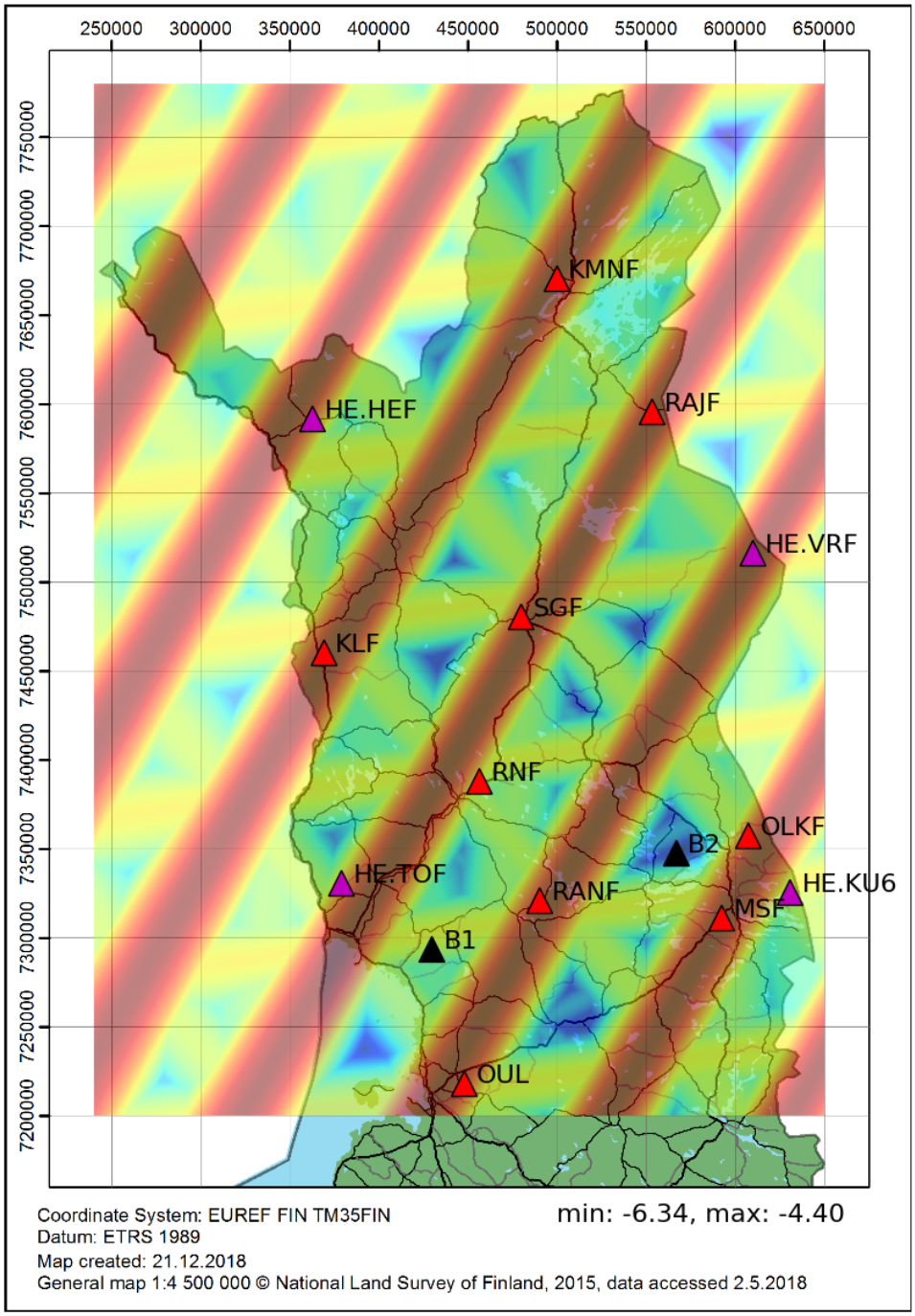


Figure 20: Minimizing sidelobes plot, the case B with 4 external stations, the second iteration. The red triangles are the FN stations, the magenta triangles are the used external stations, and the black triangles are the optimal locations for the new stations (the first and second iteration). The values (relative power) range from -6.34 (the darkest blue) to -4.40 (the darkest red).

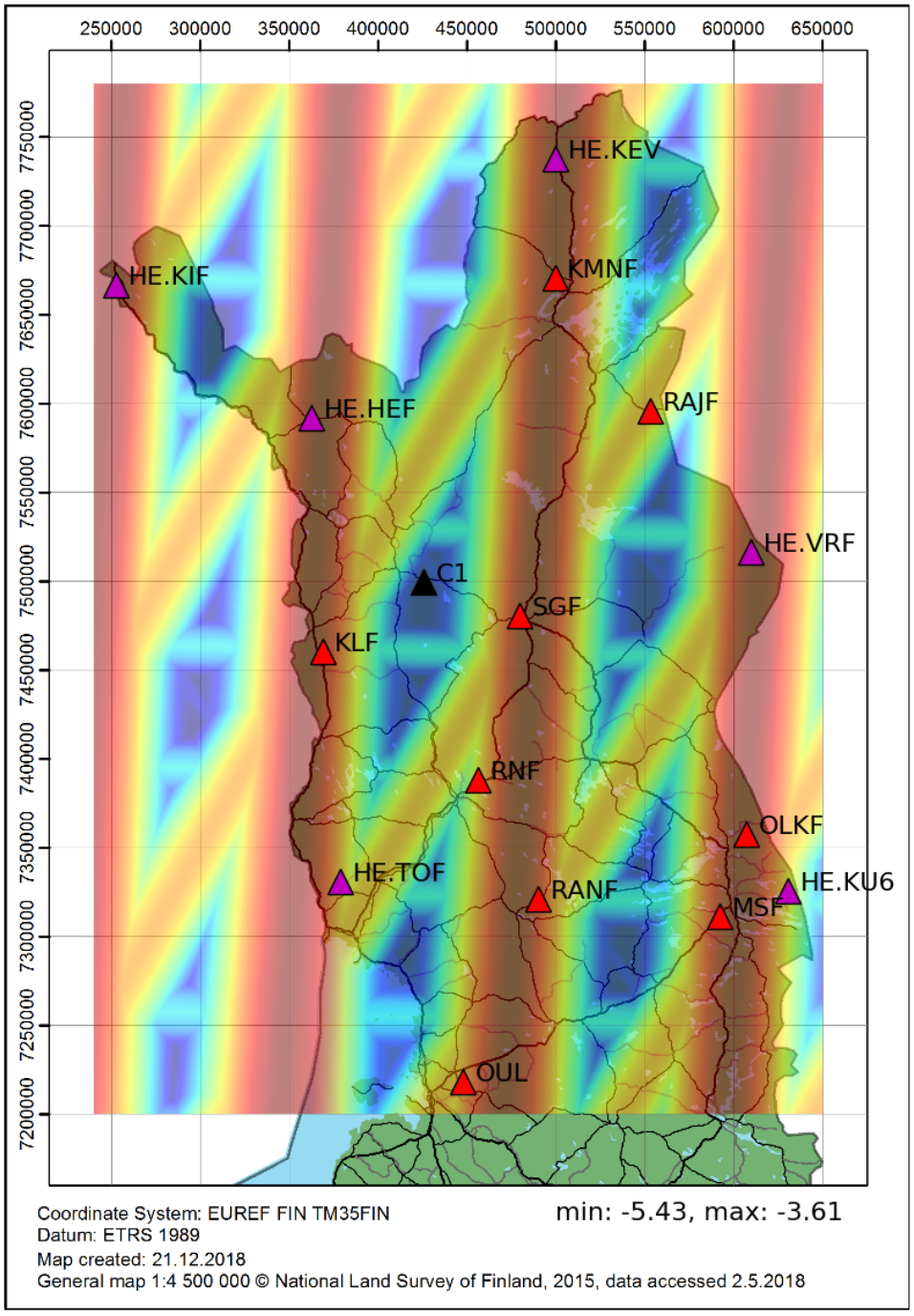


Figure 21: Minimizing sidelobes plot, the case C with 6 external stations, the first iteration. The red triangles are the FN stations, the magenta triangles are the used external stations, and the black triangle is the optimal location for the new station. The values (relative power) range from -5.43 (the darkest blue) to -3.61 (the darkest red).

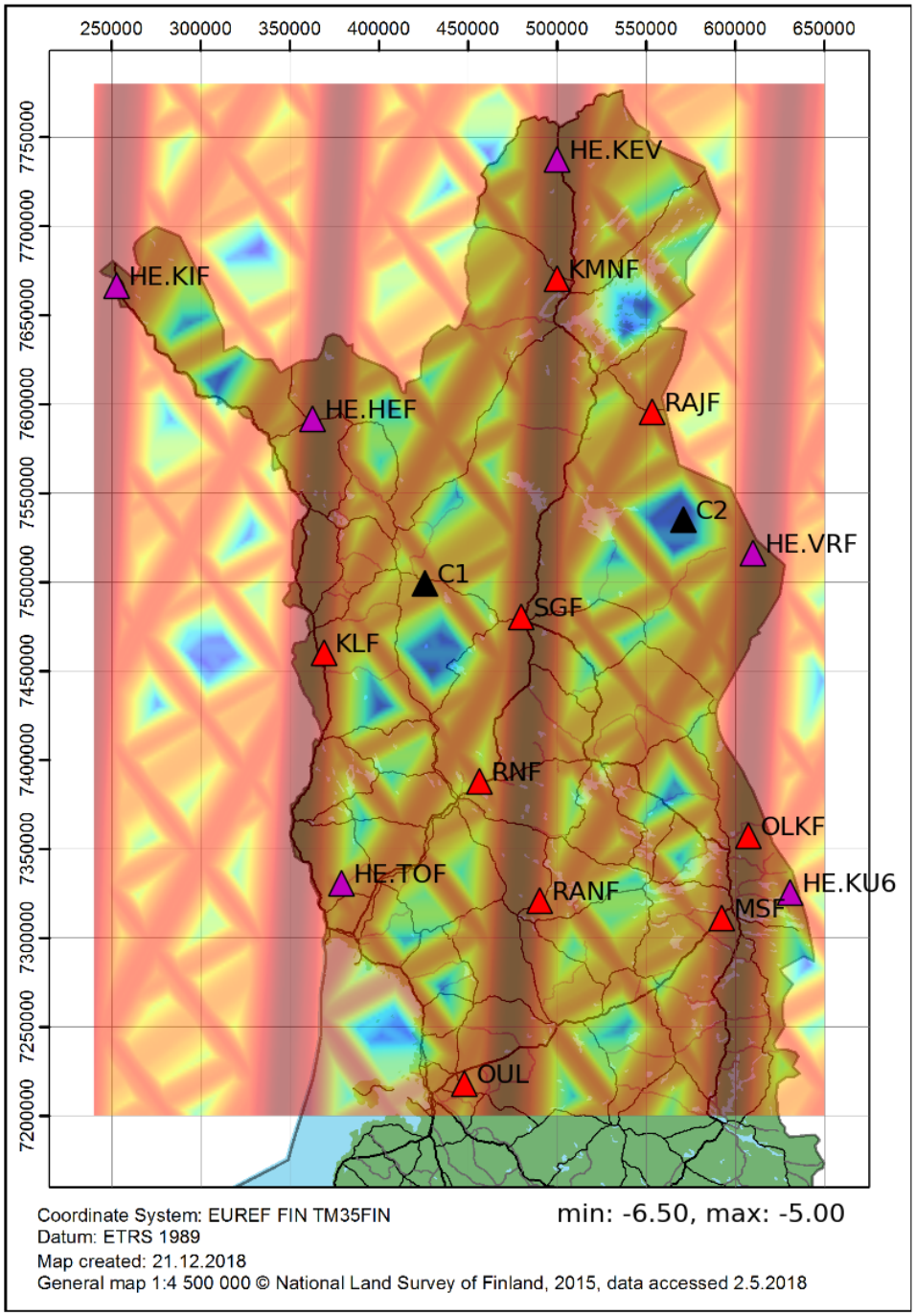


Figure 22: Minimizing sidelobes plot, the case C with 6 external stations, the second iteration. The red triangles are the FN stations, the magenta triangles are the used external stations, and the black triangles are the optimal locations for the new stations (the first and second iteration). The values (relative power) range from -6.50 (the darkest blue) to -5.00 (the darkest red).

5 Discussion

As it was made clear, only surface waves with pretty narrow frequency band were studied in this thesis. For example, P and S waves were untouched which means that detection of body waves could be one of the future interests. And as in Chapter 4.1.1 was mentioned, some body wave phases were probably registered. This suggests that the array can be potentially useful for detecting long-period body waves of strong seismic events.

Since the FN network consists of 3-component sensors and since only vertical components were used in this analysis, it is obvious that some possibilities were missed out and the the potential of the array was not fully accessed in this study. 3-component array makes it possible to analyze in more detail particle motions, i.e., polarizations of the seismic phases, which for example could have been used to distinguish between Love and Rayleigh waves and other seismic phases simply by observing the different polarizations of the waves.

As it is possible to resolve the BAZ using recording of single 3-component station, application of this kind of localization alongside array methods might be worthwhile. Note that this method leaves a 180° ambiguity to the BAZ values. (Bormann and Siegfried, 2012)

Methods and possibilities concerning 3-component arrays are discussed in detail for example by Jurkevics (1988).

As for earthquake location tests, in this study, only BAZ of the event was calculated. In order to get at least the distance to the epicenter additional information is needed. One method to solve the distance is to refer to the seismic phase travel time curves, for example, tables by Kennett and Engdahl (1991) and Jeffreys and Bullen (1940, 1948, 1958, 1967, 1970). The distance estimate is achieved by matching the observed time

difference between certain seismic phases to the reference values described in the above tables. As a side note, the SNR gain achieved through beamforming will provide more seismic phases for closer investigation.

Another way to locate the epicenter completely, and probably achieve better localization results in general, is to use multiple arrays, a minimum of two arrays. With each array, BAZ of the event is resolved, and the epicenter can be found as a crossing of great circles passing through the central station of each array.

In addition, larger catalogs with more earthquakes and longer time period, would be desirable to acquire more comprehensive earthquake localization test results. Three of the latest FN network stations started to provide test data at the end of 2017, and this determines the earthquake catalog starttimes in this thesis.

One improvement which would also increase the amount of data could be utilizing more sophisticated trigger methods to pick up lower magnitude events and specific seismic phases. This could be especially useful in case of arctic events since generally lower magnitudes of those events caused problems for detection. The advanced triggers could also make use of 3-component stations by observing different polarizations.

Another improvement concerning triggering would be multi-station triggering where the phase is registered in multiple stations. The triggered phase has to be coherent in each station. In this study, only one station at a time (SGF or RAJF) was used for triggering.

The new seismic site searching method based on array processing (Chapter 4.3) could be improved by adding restrictions to the search in order to discard completely unsuitable site locations at the earliest stages of site selection process. At the same time, it is possible to assign weights to the most eligible locations to reveal the locations meeting most of the criteria. These restrictions could include for example information about nearby roads, power lines, transformers, and other prebuilt infrastructure. Along with

positive features in the vicinity, sources of seismic noise known beforehand could be included. These additions would further highlight the suitability of possible locations for a new seismic site.

6 Conclusions

In the thesis, the properties of the FN network as an array were discussed, and the most suitable frequency and slowness combinations of seismic phases were studied. With applicable slowness and frequency combinations and with phases originating from large epicentral distances, the FN network can be utilized as a seismic array.

In Chapter 4.1, FN network was used as an array to resolve BAZ directing to the earthquake epicenter. Beampacking method was used to estimate BAZ and apparent slowness of the seismic phases. The localization tests were relatively successful and demonstrated the array processing capabilities of the FN network, although localization of events with epicenters too close to the array or with low magnitudes turned out to be problematic. In the localization tests, surface wave phases were used. The used seismic data was 15 mHz to 25 mHz bandpass filtered vertical component data from the FN network stations.

If beampacking method is used successfully to resolve the slowness and BAZ of the seismic phase, the results can be used for beamforming to increase SNR. The increase in SNR using the FN network as an array was achieved. With the frequency band in question, signal coherency of the used seismic phases was good, and the seismic noise was relative incoherent across the array. Coherent signals and incoherent noise are the main requirements for successful beamforming.

In Chapter 4.3, a method utilizing array processing to search a location for a new seismic site is presented. In this method, an initial station configuration is defined. The method

is looking for a new site location that reduces the sidelobes in the array transfer function the most. The results can be used as an additional consideration when planning a new site location. Also, the visualization of the array geometry in the resulting sidelobe minimizing plots might be of use. The presented method solely tries to minimize the ambiguities in the detection of seismic phases with certain wavenumber vectors. And thus, broaden the detectable wavenumber passband described in Chapter 3.3.

References

Aki, K. and Richards, P. G. *Quantitative Seismology*. University Science Books, 2nd edition, 2002. ISBN 0-935702-96-2.

Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., and Wassermann, J. Obspy: A Python Toolbox for seismology. In *Seismological Research Letters*, volume 81, page 530, 2010. doi: 10.1785/gssrl.81.3.530. URL `http://dx.doi.org/10.1785/gssrl.81.3.530`.

Bormann, P. and Siegfried, W. Earthquake location at teleseismic distances from 3-component records (Tutorial with exercise by hand). In Bormann, P., editor, *New Manual of Seismological Observatory Practice 2 (NMSOP-2)*, pages 1 – 18. Deutsches GeoForschungsZentrum GFZ, Potsdam, 2012. doi: {10.2312/GFZ.NMSOP-2_EX_11.2}.

Eloranta, E. *Geofysiikan kenttäteoria*. Säteilyturvakeskus, 2003. ISBN 978-952-478-194-7. URL `http://urn.fi/URN:NBN:fi-fe2014120249023`.

EPOS. European Plate Observing System, n.d. URL `https://www.epos-ip.org/`. Accessed: 14.2.2019.

Finnish Geodetic Institute. Coordinate Transformation Service, 2007. URL `http://coordtrans.fgi.fi/transform.jsp`. Accessed: 20.7.2018.

Gillies, S. shapely, geometric objects, predicates, and operations, 2018. URL `https://pypi.org/project/Shapely/1.6.4.post1/`. Accessed: 15.2.2019.

Gillies, S. et al. The Shapely User Manual, 2018. URL `https://shapely.readthedocs.io/en/latest/manual.html`. Accessed: 15.2.2019.

Grotzinger, J. P. and Jordan, T. H. *Understanding Earth: Seventh Edition*. Macmillan Learning, 7th edition, 2014. ISBN 9781464138744.

Havskov, J., Ottemöller, L., Trnkoczy, A., and Bormann, P. Seismic Networks. In Bormann, P., editor, *New Manual of Seismological Observatory Practice 2 (NMSOP-2)*, volume 2, pages 1 – 65. Deutsches GeoForschungsZentrum GFZ, Potsdam, 2012. doi: {10.2312/GFZ.NMSOP-2_ch8}.

Hunter, J. D. Matplotlib: A 2D graphics environment. In *Computing In Science & Engineering*, volume 9, pages 90–95. IEEE COMPUTER SOC, 2007. doi: 10.1109/MCSE.2007.55.

International Seismological Centre. *The International Seismograph Station Registry*. Internatl. Seismol. Cent., Thatcham, United Kingdom, 2015. URL `http://www.isc.ac.uk/registries/`. Accessed: 29.6.2018.

International Seismological Centre. *ISC Bulletin*. Internatl. Seismol. Cent., Thatcham, United Kingdom, 2018. URL `http://www.isc.ac.uk/iscbulletin/search/catalogue`. Accessed: 22.10.2018.

Jeffreys, H. and Bullen, K. E. Seismological tables. page 50. London: British Association for the Advancement of Science, Gray Milne Trust, London, 1940, 1948, 1958, 1967, 1970.

Johnson, D. and Dudgeon, D. *Array Signal Processing: Concepts and Techniques*. Prentice-Hall signal processing series. P T R Prentice Hall, 1993. ISBN 9780130485137.

Jones, E., Oliphant, T., Peterson, P., et al. SciPy: Open source scientific tools for Python, 2001–. URL `http://www.scipy.org/`. Accessed: 15.2.2019.

Jurkevics, A. Polarization analysis of three-component array data. In *Bulletin of the Seismological Society of America*, volume 78, page 1725, 1988. URL `http://dx.doi.org/`.

Karney, C. F. F. geographiclib 1.49 documentation, 2017. URL `https://geographiclib.sourceforge.io/1.49/python/`. Accessed: 15.2.2019.

Kennett, B. and Engdahl, E. Traveltimes for global earthquake location and phase identification. In *Geophysical Journal International*, volume 105, pages 429 – 465, 05 1991.

Korja, A. and Vuorinen, T. A. Fin-epos: Finnish national initiative of the European Plate Observing System. In et. al., K. I., editor, *Lithosphere 2016 : Ninth Symposium on the Structure, Composition and Evolution of the Lithosphere in Finland : Programme and Extended Abstracts*, volume 65 of *University of Helsinki, Institute of Seismology, Report-S*, pages 63–64. University of Helsinki, Institute of Seismology, 2016. ISBN 978-952-10-9282-5.

Kozlovskaya, E., Narkilahti, J., Nevalainen, J., Hurskainen, R., and Silvennoinen, H. Seismic observations at the Sodankylä Geophysical Observatory: history, present, and the future. In *Geoscientific Instrumentation, Methods and Data Systems*, volume 5, pages 365–382, 08 2016.

Krischer, L., Megies, T., Barsch, R., Beyreuther, M., Lecocq, T., Caudron, C., and Wassermann, J. ObsPy: a bridge for seismology into the scientific Python ecosystem. In *Computational Science & Discovery*, volume 8, page 014003, 2015. URL `http://stacks.iop.org/1749-4699/8/i=1/a=014003`.

Lay, T. and Wallace, T. C. *Modern Global Seismology*. Number v. 58 in International Geophysics Series. Academic Press, 1995. ISBN 9780127328706. URL `http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=207438&site=ehost-live`.

Matplotlib Developers. matplotlib: v1.5.3 (version v1.5.3). Zenedo, 2016. doi: http://doi.org/10.5281/zenodo.61948.

McNamara, D. and Buland, R. P. Ambient noise levels in the continental united states. In *Bulletin of the Seismological Society of America*, volume 94, pages 1517–1527, 09 2004.

Megies, T., Beyreuther, M., Barsch, R., Krischer, L., and Wassermann, J. ObsPy – What can it do for data centers and observatories? In *Annals of Geophysics*, volume 54, pages 47–58, 2011. doi: 10.4401/ag-4838. URL `https://www.annalsofgeophysics.eu/index.php/annals/article/view/4838`.

Mykkeltveit, S., Ringdal, F., Kvaerna, T., and W. Alewine, R. Application of regional arrays in seismic verification. In *Report, 1 May - 31 Jul. 1990 Royal Norwegian Council for Scientific and Industrial Research, Kjeller.*, volume 80, 08 1990.

Narkilahti, J., Kozlovskaya, E., Silvennoinen, H., Hurskainen, R., and Nevalainen, J. Upgrading the Northern Finland Seismological Network. In *EGU General Assembly Conference Abstracts*, volume 18 of *EGU General Assembly Conference Abstracts*, pages EPSC2016–10224, April 2016.

NORSAR. Large nuclear test in North Korea on 3 September 2017, 2017. URL `https://www.norsar.no/press/latest-press-release/archive/large-nuclear-test-in-north-korea-on-3-september-2017-article1534-984.html`. Accessed: 2.8.2018.

Oliphant, T. E. NumPy: A guide to NumPy. USA: Trelgol Publishing, 2006–. URL `http://www.numpy.org/`. Accessed: 15.2.2019.

Olsen, K. G. and Nettles, M. Patterns in glacial-earthquake activity around Greenland, 2011–13. volume 63, page 1077–1089. Cambridge University Press, 2017. doi: 10.1017/jog.2017.78.

Orfeus Data Center. Orfeus data center webdc3 web interface, 2018. URL `https://orfeus-eu.org/webdc3/`. Accessed: 10.12.2018.

Peterson, J. et al. Observations and modeling of seismic background noise. U.S. Geological Survey open-file report. US Geological Survey Albuquerque^ eNM NM, 1993.

Rost, S. and Thomas, C. Array Seismology: methods and Applications. In *Reviews of Geophysics*, volume 40, page 1008, 12 2002.

Rost, S. and Thomas, C. Improving Seismic Resolution Through Array Processing Techniques. In *Surveys in Geophysics*, volume 30, pages 271–299, Oct 2009. doi: 10.1007/s10712-009-9070-6. URL `https://doi.org/10.1007/s10712-009-9070-6`.

Schweitzer, J., Fyen, J., Mykkeltveit, S., Gibbons, S. J., Pirli, M., Kühn, D., and Kværna, T. Seismic Arrays. In Bormann, P., editor, *New Manual of Seismological Observatory Practice 2 (NMSOP-2)*, volume 2, pages 1 – 80. Deutsches GeoForschungsZentrum GFZ, Potsdam, 2012. doi: {10.2312/GFZ.NMSOP-2_ch9}.

Shearer, P. M. *Introduction to Seismology*. Cambridge University Press, 1999. ISBN 9780521669535.

The Global CMT Project. Global CMT Web Page, 2018. URL `https://www.globalcmt.org/`. Accessed: 8.3.2019.

The ObsPy Development Team. ObsPy: A Python Toolbox for seismology/seismological observatories. 2017. doi: 10.5281/zenodo.165135. URL `https://docs.obspy.org/`.

Trnkoczy, A. Understanding and parameter setting of STA/LTA trigger algorithm. In Bormann, P., editor, *New Manual of Seismological Observatory Practice 2 (NMSOP-2)*, pages 1 – 20. Deutsches GeoForschungsZentrum GFZ, Potsdam, 2012. doi: {0.2312/GFZ.NMSOP-2_IS_8.1}.

Wang, J. A Scheme for Initial Beam Deployment for the International Monitoring System Arrays. In *pure and applied geophysics*, volume 159, pages 1005–1020, Mar 2002. doi: 10.1007/s00024-002-8670-6. URL `https://doi.org/10.1007/s00024-002-8670-6`.

Whitaker, J. pyproj, Python interface to PROJ.4 library, 2016. URL `https://pypi.org/project/pyproj/1.9.5.1/`. Accessed: 15.2.2019.

Appendix: Of Codes and Scripts

The acquisition of the results in the thesis included plenty of coding and utilization of ready-made computer scripts. As the main programming tool, Python (version 3.5.3) language was used.

For the thesis, a collection of array processing scripts was made, taking advantage of many external packages already made for Python by numerous contributors. In this appendix, some of the used scripts and functions from the packages are reviewed.

For data manipulation and mathematical purposes, SciPy (version 0.19.1) and NumPy (version 1.13.1) packages were used. (Jones et al., 2001–; Oliphant, 2006–)

Another notable external package used was ObsPy framework (version 1.1.0). (The ObsPy Development Team, 2017; Beyreuther et al., 2010; Megies et al., 2011; Krischer et al., 2015)

All plotting and creation of figures was done using ObsPy and matplotlib library (version 1.5.3). (Matplotlib Developers, 2016; Hunter, 2007)

Basic array methods

Using the package created for the thesis, it is possible to define an arbitrary array and apply basic array processing method using the array. The methods include plotting the station locations and calculating and plotting the transfer function and the wavenumber passband. The wavenumber basspand methods are based on calculations presented by Wang (2002).

The array transfer functions (Figure 8) are plotted using `array_transff_wavenumber` function (URL: `https://docs.obspy.org/packages/autogen/obspy.signal.array_analysis.array_transff_wavenumber.html`, accessed 15.2.2019) and the PPSD plots

are created using the PPSD class (URL: `https://docs.obspy.org/packages/autogen/obspy.signal.spectral_estimation.PPSD.html`, accessed: 15.2.2019), both provided in ObsPy. In addition, seismographs (Figures 4 and 5) are plotted using ObsPy.

Earthquake localization tests

Chapter 4.1 covered the earthquake localization tests carried out using FN network. For the tests, a script to automate the process and to plot and handle the results was made. As an input, the script takes in the earthquake catalog describing various parameters, such as magnitudes, origintimes, and precise locations. In addition, parameters (see Table 4) for the calculations are defined.

The script processes on earthquake at a time, as described in Chapter 4.1. The original BAZ relative to the center station (SGF) is solved using geodesic calculations provided in geographiclib (version 1.49) (Karney, 2017). The results obtained for each earthquake are stored and can be analyzed and plotted (see for example Figures 10 and 11).

While processing each earthquake, additional information about the event and the corresponding process is produced, such as STA/LTA trigger plots (Figure 5), seismographs with the beam (Figure 4), and the beampacking result plot (Figure 13). In addition, the signal cross-correlation information is stored for later use.

An additional standalone script to output noise cross-correlation data was also made. The noise cross-correlation data as well as the signal cross-correlation data produced in localization tests, can be visualized and analyzed.

Because of the relatively burdensome nature of the multiple calculations done for the localization tests, parallel processing using Python multiprocessing library was utilized.

To calculate the power in order to use beampacking method, PSD calculations were done using `psd` function from matplotlib (URL: `https://matplotlib.org/api/mlab_api.html#matplotlib.mlab.psd`, accessed: 15.2.2019).

For signal and noise cross-correlation calculations, `correlate` and `xcorr_max` functions from ObsPy were used (URL: `https://docs.obspy.org/master/packages/autogen/obspy.signal.cross_correlation.html`, accessed: 15.2.2019).

In Figures 16 and 12, the interpolation was done using `UnivariateSpline` class from SciPy (URL: `https://docs.scipy.org/doc/scipy/reference/generated/scipy.interpolate.UnivariateSpline.html` accessed: 15.2.2019). The interpolations in Figure 16 were smoothed using the smoothing factor of 0.1.

Sidelobe minimizing plots

For the sidelobe minimizing plots introduced in Chapter 4.3 an automatic location search script was made. In one run of the script, it is possible to do multiple searches for a new location with different configurations and do multiple iterations. Iteration in this matter means that the best location is included in the calculations and the next sidelobe minimizing location is searched.

There is a possibility to limit the location inside a certain country and reject locations which would increase the aperture. In this thesis, new locations are required to lie inside Finland and the original aperture must remain.

In order to change between different coordinate systems, mainly from WGS84 to ETRS_TM35FIN and vice versa, transformations provided in pyproj (version 1.9.5.1) package were used. (Whitaker, 2016)

To check the country where the coordinates are in, shapely (version 1.6.4.post1) package was used, as well as the land border information contained in `countries.geojson` file. The file was downloaded 20.12.2018 from `https://datahub.io/core/geo-countries`. (Gillies, 2018; Gillies et al., 2018)